Algorithm Theoretical Basis Document

Cloud Information Retrieval from MIPAS Measurements

Version 2.1

June 10, 2010

ABSTRACT

This document is the Algorithm Theoretical Basis Document for the Cloud Information Retrieval from MIPAS Measurements produced according to the guidelines specified as a deliverable item under the Cloud Information Retrieval from MIPAS Measurements Contract No. 20601/07/I-OL.

Prepared by:

Reinhold Spang, Karina Arndt, Anu Dudhia, Sabine Grießbach, Michael Höpfner,
Jane Hurley, John Remedios, Harjinder Sembhi, and Richard Siddans
(contact: r.spang@fz-juelich.de)

DOCUMENT STATUS SHEET

First draft	October 28, 2008
First Version	February 04, 2009
First Version extension	June 25, 2009
First Version update	September 07, 2009
Second Version	December 17, 2009
Second Version corrections	June, 2010

DOCUMENT CHANGE RECORDS
made since last issue

Version 1.0
- chapter Error and Quality Assessment added
- tbd sub-sections completed respectively deleted (formerly 2.3.2 CSDB subset)

Version 1.1
- Added section on cloud and cloud type decision algorithm

Version 1.1 Update
- Update Macro Retrieval chapter

Version 2.0
- Update data ingestion parameters
- Update cloud decision
- Update processing constraints

Version 2.1
- correction of typos

Acronyms and Abbreviations

ADP	Area Density Path (limb)
ATBD	Algorithm Technical Basis Document
BT	Brightness Temperature (here used for Blackbody Equivalent Temperature)
BTD	Brightness Temperature Difference
BTR	Blind Test Retrieval
CEF	Cloud Effective Fraction
CEX	Cloud Extinction
CI	Cloud Index
CI-A	MIPAS Cloud Index for band A operational used in the ESA L2 processing
COF	Cloud Occurrence Frequencies
CR	Colour Ratio
CSDB	Cloud Scenario Database
CTH	Cloud Top Height
CTP	Cloud Top Pressure
CTT	Cloud Top Temperature
EF	Effective Fraction of the cloud in the FOV
IWC	Ice Water Content
IWP	Ice Water Path
NAT	Nitric Acid Trihydrate
NI	NAT Index (Color Ratio)
PSC	Polar Stratospheric Cloud
PSD	Particle Size Distribution
PVP	Product Validation Procedure
R_{eff}	Effective Radius of PSD
R_{mean}	Mean Radius of PSD
SRD	Software Requirement Document
STS	Sulfuric Ternary Solutions
TH	Tangent Height
VDP	Volume Density Path (limb)

Contents

1 Introduction

1.1 Scope of the Document

The purpose of the Algorithm Theoretical Basis Document (ATBD) is to present the algorithm and technical details necessary to realise a processor for the retrieval of cloud properties form the Level 1b (L1B) data of ESA instrument MIPAS on-board the ENVISAT satellite. The infrared spectra measured by MIPAS (main information of L1B) include sufficient information to retrieve various cloud parameter. Recent publications have already shown the large potential of the dataset ([11], [12], [5], [6]). But so far no validated and consolidated MIPAS cloud product is available for the scientific community. Consequently the development of a cloud parameter processor for MIPAS and the application to the time series is highly desired.

The proposed MIPAS cloud processor follows the requirements of the statement of work of the ESA-ITT AO/1-5255/06/I-OL. Therefore the development and application of the processor is so far restricted to the first measurement period of MIPAS from July 2002 to March 2004 (spectral high resolution (HR) mode). After 2004 MIPAS is operating with slightly reduced resolution (RR-mode). A transfer of the prototype processor for application of the RR mode has been taken into account during the algorithm development. It is desirable, that this will be also considered in the realisation phase of the prototype processor.

Primary task of the software is to retrieve cloud properties from the MIPAS L1B spectra. This contains various items of interest:

- the detection of cloudy spectra in the L1B data
- to classify various cloud types in the measurements (e.g. polar stratospheric clouds, liquids and ice clouds)
- to retrieve cloud top information on height, temperature, pressure
- to retrieve profile information on cloud parameters (e.g.. extinction, ice water path or integrated limb path quantities like volume/area density path)
- information on microphysical parameter like effective radius of the particle size distribution (PSD) or Volume and Area densities.

The retrievability of the cloud parameters like the ones above have been investigated in a feasibility study of the MIPclouds project and are summarised in the Technical Note [1](in the following abbreviated with TN).

1.2 Outline of the Processor

For the better understanding of the proposed algorithms in the following sections it is necessary to have a brief overview and understanding about the proposed sequence of processing steps and the complete flowchart of the prototype processor. Despite the detailed processor flowchart is designed in a later stage of the project, a draft of the scientific processor flowchart (based on the proposal of the project and on results of the TN) is already designed and described in the following subsections.

Figure 1.1 gives an overview of the outlined processor in conjunction with scientific analysis. The outcome of scientific analysis is closely linked with potential retrieval parameter of the processor. The Cloud Scenario Database (CSDB) was crucial for the algorithm development and the error assessment of the parameters. The database contains more than 70000 different cloud scenarios and close to 600000 cloud spectra (for details s. [1]). In the processor only a subset of CSDB spectra – condensed information like mean spectra – are intended to be used.

The main processing steps outlined in the rough processing flowchart of Figure 1.1 are:

- Detection
- Classification
- Macrophysical Retrieval
- Microphysical Retrieval

At each of this steps various input parameters and look-up tables are necessary, respectively various output parameters will be created. This is described in more detail in section 1.2.2. Together with the four items above this is describing the main sections of the ATBD.

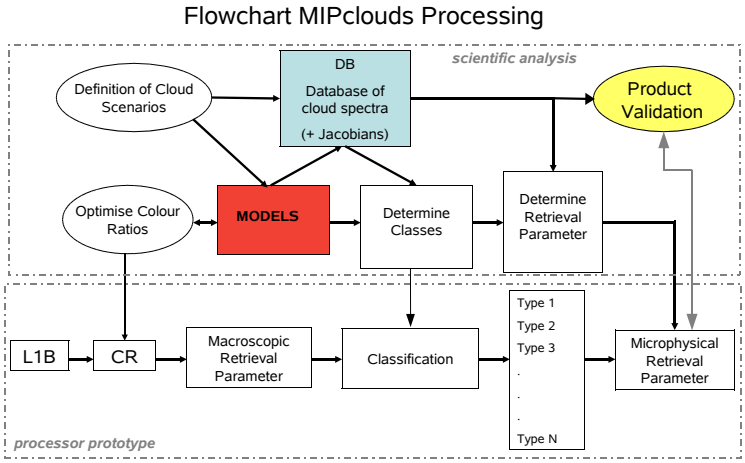

Figure 1.1: Overview of the MIPclouds processing scheme in interaction with the scientific analysis and validation activities

1.2.1 Parameter of Interest

Various retrieval parameter have been investigated in the TN [1]. A summary of the potential parameter is given in Table 1.1. For each of the listed parameter it is necessary to define an adequate validation dataset in respect of data quality, error budget and statistical significance (e.g. number of useful coincidences) of the validation dataset [2]. Retrieval schemes for most of the parameter will be addressed in the following section. The development of the prototype processor has the strong constrain, that the processing shall work in near real time (NRT). This excludes any retrieval scheme which handles the complex and time consuming radiative transport in the presence of clouds (e.g. including single or multiple scattering). Simplifications and parameterisation based on the CSDB have been developed to setup a fast processing scheme for cloud parameter retrievals.

Parameter	**Symbol**	comments on **validation method**, errors, etc.
cloudiness flag	cf	status flag for each spectrum 1/0
Cloud occurrence frequencies	COF	statistical means: SM (e.g. zonally, seasonally)
Cloud Top Height	CTH	**SM, BTR**
Cloud Top Temperature	CTT	**SM, BTR**
Cloud Top Pressure	CTP	**SM, BTR**
Cloud Base Height	CBH	**SM, BTR**
Cloud Extinction	CEX	**SM, CM**
Cloud Effective Fraction	CEF	Consideration of FOV effects $\Rightarrow$ better CTH
Cloud Classification		
Stratospheric Cloud Types:		**CM, BTR** for all types
- Nitric Acid Trihydrate	NAT	
- Sulfuric Ternary Solutions	STS	
- Ice	ICE	
- Aerosol	AER	
Cloud types in the		
free troposphere and UTLS:		**CM, BTR** for all types
- Cirrus / Ice clouds	CirC	
- Liquid	LiqC	
- Aerosol	AerC	
Area Density Path	ADP	**BTR**; along the limb path; for threshold values and estimates only;
(1) limb Ice/Liquid Water Path respectively Vol. Density Path	IWP / LWP VDP	**BTR, (CM)**; quantities along the limb path
(1) limb Ice/Liquid Water Content and Volume Density respectively	IWC / LWC VD	**BTR, SM, (CM)**;only estimate available
Effective and/or mean radius	R_{eff} / R_{mean}	**BTR, SM, (CM)**;only course size bin retrieval

Table 1.1: MIPAS cloud processor output parameter with short cut comments for **SM**: statistical means (zonally or seasonally means for various altitude bins), **CM**: coincidence method for the validation dataset, **BTR**: blind test retrievals. For details on validation method see [2].
(1) For transformation from limb path or slant column quantities like IWP to nadir IWP or respectively IWC one has to assume the simplification of a homogeneous cloud layer filling the complete tangent height layer.

1.2.2 Intended Processing Scheme

A more detailed outline of the intended Processing scheme is summarised in Figure 1.2 and Figure 1.3. Various input data sets are necessary for the processing and are described in section 2. Single steps of the processing are described in the following section. After merging various data sources (L1a, L2, ECMWF data) a preprocessing takes place to create a consolidated profile based dataset of radiances for a number of specified micro windows. The cloud detection is then a crucial part of the processor where various methods can be applied and add-on information like the influence of high water vapour continuum needs to be retrieved for the best possible decision which spectra of a profile are influenced by cloud emissions (Fig. 1.2). After this step the retrieval of cloud properties is starting (Fig.1.3)

It is recommended to compile two output datasets. One set with detailed information on the processing steps and the various detection methods and a second condensed product output which includes only the main target parameter for the scientific users.

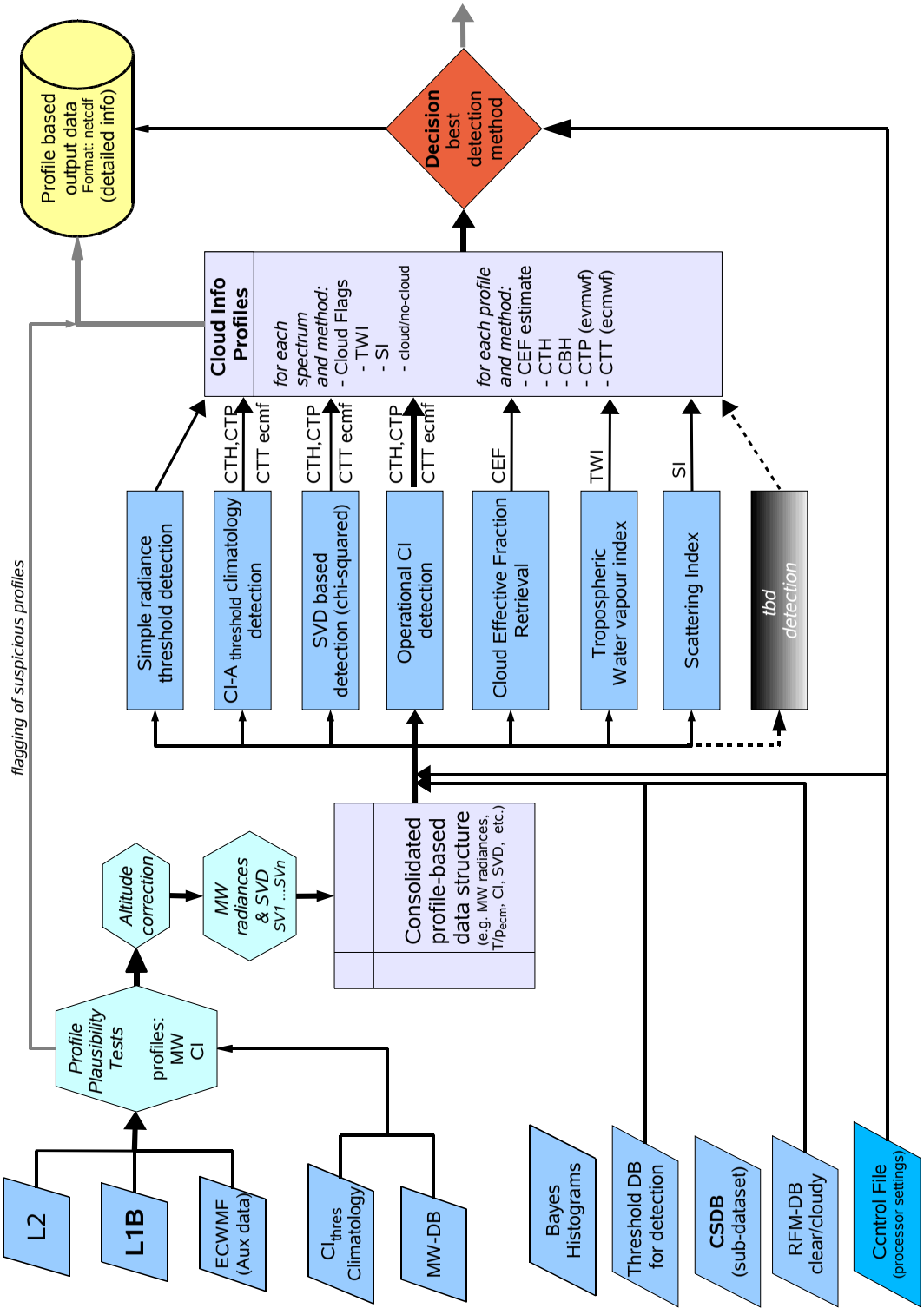

Figure 1.2: Sketch for the prototype processor flowchart including proposed input and output files and processing steps **Part 1**

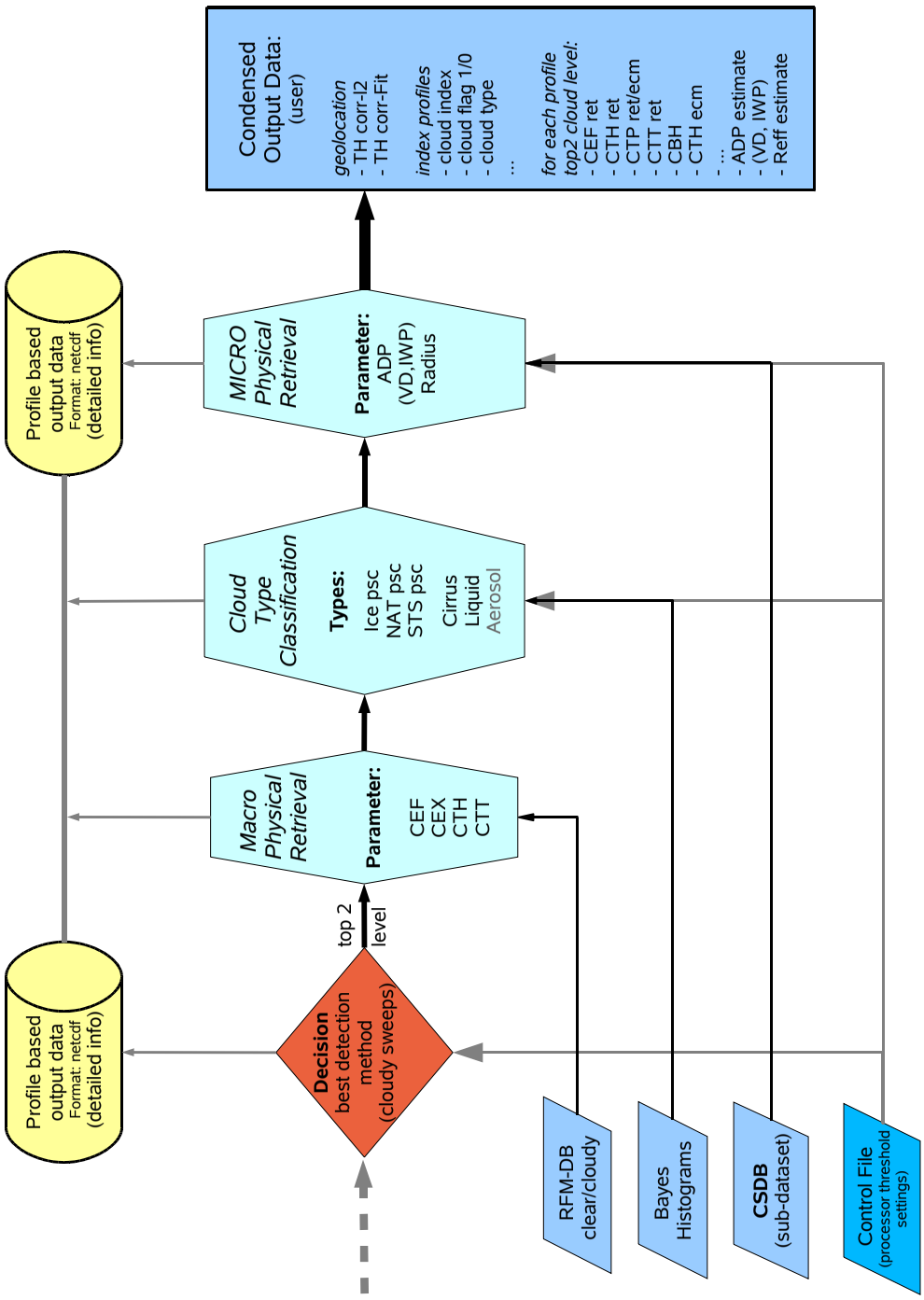

Figure 1.3: Sketch for the prototype processor flowchart including proposed input and output files and processing steps **Part 2**

2 Data Ingestion

2.1 MIPAS L1B/L2 data

Level 1B data for the ESA processing are the prime input for the processing, whereby the IR spectra are the main data volume. Additional parameters of the L1B file are necessary for the processing steps and for characterisation of the produced output files (s. Table 2.1)

Symbol	**Unit**([2])	**Description**
TIME		MIPAS modified Julian Date (ref. JD2000)
Xsc, Ysc, Zsc	km	S/C position vector in earth-fixed reference
LOSazim	degrees	Line of sight azimuth
LOSelev	degrees	Line of sight elevation
TH	km	Altitude
lat	degrees	Geo Coordinate Latitude
lon	degrees	Geo Coordinate Longitude
REarth	km	Radius of earth surface curvature in looking direction at nadir of LOS tangent point
rad	W/cm^2 sr cm^{-1}	Mipas L1B spectra
qual_flag		Quality indicator (PCD), 0 = non-corrupted, 1 = one or more bands corrupted, -1 = blank MDSR
attach_flag		Attachment Flag (set to 1 if all MDSRs corresponding to this ADSR are blank, set to zero otherwise)
retrieval_p_t_flag		Flag indicating successful p,T retrieval
retrieval_vmr_flag		Flag indicating successful VMR retrievals

Table 2.1: MIPAS L1B parameter of interest

Retrieval results of the level 2 processing can be a valuable information for various processing steps. So far only a very limited selection is envisaged for the cloud processor (s. Table 2.2).

Symbol	**Unit**	**Description**
TIME		MIPAS modified Julian Date (ref. JD2000)
Temp	K	Temperature
Press	hPa	Pressure
Alt	km	Altitude

Table 2.2: MIPAS L2 parameter of interest

[2] original input parameter of L1B might need conversion to the required unit specified in the table

2.2 ECMWF data

Table 2.3 shows the parameter of interest of the ECMWF data base. The parameter will be used for a priori information, for a planned altitude correction and for temperature information, like supposed in some of the following algorithms. For each MIPAS L1B/L2 profile a coincident ECMWF profile of the parameters T, p, RH and Z needs to retrieved. This allows for example to address the first cloudy sweep in a profile and ECMWF temperature from cloud top temperature (CTT).

Symbol	Unit	Description
lon	degrees	longitude
lat	degrees	latitude
plev	hPa	isobaric pressure levels
Z	m^2/ s^2	geopotential
T	K	temperature
RH	%	relative humidity

Table 2.3: ECMWF Aux parameter of interest

2.3 Look-up tables

2.3.1 RFM Simulations

The Reference Forward Model (RFM) is a GENLN2-based line-by-line originally developed by Dudhia at Atmospheric, Oceanic and Planetary Physics, Oxford University which provides reference spectral calculations for MIPAS [13]. The RFM has been used to simulate sections of the MIPAS spectra which are either directly used in, or used in the development of, techniques utilised in this project.

2.3.1.1 For Use in Macrophysical Parameter Retrieval

The Macrophysical Parameter Retrieval, discussed in Chapter 5, requires a gas correction be applied to real MIPAS spectra before retrieving. To this end, the RFM has been used to simulate radiance spectra in the spectral range of 960 – 961 cm^{-1} at altitudes of 6 km, 9 km, 12 km, 15 km, 18 km, 21 km corresponding to standard reference atmospheres representing tropical, mid-latitudinal day and night, polar summer and polar winter conditions.

2.3.1.2 For Use in SVD Detection Methods

In the development of SVD methods (discussed in Section 3.5), as well as in the optimisation of the location of CI microwindows (Section 3.2), an ensemble of clear and cloudy spectra were compiled using the RFM. These spectra include all combinations of altitudes, extinction coefficients, cloud top heights and atmospheric variations as discussed in Table 2.4. In total, the ensemble has 5184 different atmospheric conditions: 576 of

TH [km]	k_{ext} *[km^{-1}]*	*Reference Atmosphere*	*CTH relative to TH [km]*
6, 9, 12, 15, 18, 21	0.001, 0.01, 0.1	standard mid-latitudinal, tropical, polar summer and polar winter reference atmospheres, their one standard-deviation variants, and separate perturbations in temperature, pressure, water vapour and ozone of each	-2.0, -1.5, -1.0, -0.5, 0.0, 0.5, 1.0, 1.5, 2.0

Table 2.4: Parameters used to create ensemble of cloudy atmospheres. Reference atmospheres compiled by Remedios (2001).

which are totally clear (i.e.. cloud top height $= -2.0$ km with respect to the tangent height at the centre of each MIPAS FOV) and 4608 of which contain some finite amount of cloud in the corresponding MIPAS FOV (here, when cloud top height > -2.0 km). These simulations have been carried out at the MIPAS full-resolution of 0.025 cm^{-1}.

2.3.2 Micro Window Database and its applications to L1B

Various retrieval schemes are optimised to different wavelength region. In general only small wavenumber regions – so-called micro windows (MW) – of the MIPAS spectrum are used for detection, classification or micro/macro parameter retrieval. A list of minimum and maximum wavenumber will be prepared as input data set. The MW radiance will be compiled by a simple mean over the single radiances in the MW. For the L1B of the MIPAS HR-mode these are ($\Delta\nu = 0.025$) 40 radiance per wave number. But the wavenumber windows do not necessarily have to be continuous. Certain regions in the MW range can be flagged, for example to exclude single line emissions.

For further analyses it is necessary to compute the black body equivalent temperature for the mean radiances. This quantity is often described with brightness temperature (BT). For the calculation it is necessary to invert the Planck function, like described in subsection 2.3.2.2. But for the calculation of mean radiances it will be necessary to apply an apodization to the L1B spectra.

2.3.2.1 Apodization of L1B spectra

podization is used to reduce spurious spectral features in the instrument line-shape of a Fourier transform spectrometer. The reduction of side-lobes is achieved at the expense of a slight reduction in spectral resolution. The unapodized Envisat MIPAS spectra provided by ESA are apodized by means of the strong Norton-Beer apodization. This apodization provides the best compromise in terms of degradation in resolution, reduction of side-lobes, convergence behaviour of side-lobes and correlations introduced in the spectral grid points. For apodization the unapodized MIPAS spectra are directly convoluted with the strong Norton-Beer apodized instrument line shape, i. e.

$$f'(\nu) = \sum_{i=0,2,4} C_i \, Q_i(2\pi\nu\Delta), \tag{2.1}$$

$$C_0 = 0.045335, \quad C_2 = 0.554883, \quad C_4 = 0.399782, \tag{2.2}$$

$$Q_0(a) = \operatorname{sinc} a = sin(a)/a, \tag{2.3}$$

$$Q_2(a) = -\frac{8}{a2}\left[\left(1-\frac{3}{a2}\right)\operatorname{sinc} a + \frac{3}{a2}\cos a\right], \tag{2.4}$$

$$Q_4(a) = \frac{384}{a4}\left[\left(1-\frac{45}{a2}+\frac{105}{a4}\right)\operatorname{sinc} a + \frac{5}{a2}\left(2-\frac{21}{a2}\right)\cos a\right]. \tag{2.5}$$

The argument is $a = 2\pi\nu\Delta$ with the maximum optical path difference Δ. Series expressions are required to evaluate the functions Q_i for small arguments to avoid numerical singularity of the sinc function. For efficient numerical computation the apodized instrument line shape will be calculated once at program start and stored as a look-up table.

2.3.2.2 Planck radiation law and Brightness Temperature

Plancks law of black body radiation describes the radiance from a black body emitted per unit area and solid angle at a specific frequency, wave length or wave number and temperature. Here the wave number convention is used:

$$I(\tilde{\nu},T) = \mathrm{c}_1\tilde{\nu}^3\frac{1\cdot 10^4}{exp(\frac{\mathrm{c}_2\tilde{\nu}}{T})-1} \tag{2.6}$$

$$\mathrm{c}_1 = 2\mathrm{hc}^2 \qquad \text{in } \mathrm{Wm}^2 \tag{2.7}$$

$$\mathrm{c}_2 = \frac{\mathrm{hc}}{\mathrm{k}}\cdot 100 \qquad \text{in cm K} \tag{2.8}$$

with:
$I(\tilde{\nu},T)$ - radiance in $\frac{W}{cm^2\ sr\ cm^{-1}}$
$\tilde{\nu}$ - wave number in cm^{-1}
T - Temperature in K
c_1, c_2 - radiation constants
h = $6.62600755\cdot 10^{-34}\ Js^2$ Planck constant
c = $2.99792458\cdot 10^8\frac{m}{s}$ speed of light in vacuum
k = $1.380658\cdot 10^{-23}\frac{J}{K}$ Boltzmann constant

The temperature a black body is supposed to have to emit the same radiance as measured is called the brightness temperature. This assumption holds only for thermal equilibrium.

Solving equation (2.6) for T yields the brightness temperature:

$$BT = \frac{c_2\tilde{\nu}}{ln(\frac{c_1\tilde{\nu}^3\cdot 10^4}{I(\tilde{\nu})}+1)} \tag{2.9}$$

The last formula has to be applied for the calculation of BT – better described with equivalent blackbody temperature – for the corresponding selection of MWs and respective mean radiance profiles. Despite formula (2.9) is only valid for a fixed wavenumber and not for the mean radiance of the complete MW, this simplification is sufficient for the follow-on application. The Brightness Temperature Difference (BTD) is then simply the difference of two selected MWs at the mean wavenumber of the corresponding MWs:

$$BTD(MW_1,MW_2) = BT(r_1,MW_1) - BT(r_2,MW_2) \tag{2.10}$$

with $r_{1,2}$ = mean radiances in $W/cm^2\ sr\ cm^{-1}$ for $MW_{1,2}$ and $MW_{1,2}$ = mean wavenumber of $MW_{1,2}$.

2.3.3 Bayes Histograms

For a statistical classification of different cloud types a "naive Bayes classifier" has been applied to CSDB spectra. This is a simple probabilistic classifier based on applying Bayes' theorem with strong (naive) independence assumptions. A more descriptive term for the underlying probability model would be "independent feature model". The classifier is trained by utilising the cloud radiance database prepared from the CSDB (one wavenumber resolution subset of spectra). Individual MIPAS measurements are classified based on colour ratios and brightness temperature differences derived from the corresponding radiance data. In spite of the over-simplified assumptions, the naive Bayes classifier seems to work reasonably well in most cases. Results are extremely sensitive to size of the bins used for the probability histograms.

To give an overview of the method the following description from [22] is taken to hand.
Naive Bayes classifiers can handle an arbitrary number of independent variables whether continuous or categorical. Given a set of variables, $X = \{x_1, x_2, \ldots, x_d\}$, the aim is to construct the probability for the event C_j among a set of possible outcomes $C = \{c_1, c_2, \ldots, c_d\}$. In a more familiar language, X is the predictors and C is the set of categorical levels present in the dependent variable. Using Bayes' rule:

$$p(C_j|x_1, x_2, \ldots, x_d) = p(x_1, x_2, \ldots, x_d|C_j)p(C_j) \tag{2.11}$$

where $p(C_j|x_1, x_2, \ldots, x_d)$ is the posterior probability of class membership, i.e., the probability that X belongs to C_j. Since Naive Bayes assumes that the conditional probabilities of the independent variables are statistically independent we can decompose to a product of terms

$$p(X|C_j) = \prod_{k=1}^{d} p(x_k|C_j) \tag{2.12}$$

and rewrite the posterior as

$$p(C_j|X) = p(C_j) \prod_{k=1}^{d} p(x_k|C_j) \tag{2.13}$$

Using Bayes' rule above, we label a new case X with a class level C_j that achieves the highest posterior probability.

The training of the classifier:

For all input classes of spectra that represent a certain cloud class all possible colour ratios and brightness temperatures are computed. The results are subsumed to histograms for each class and ratio. The next step is to identify the ratios with the highest information content for cloud classification. Therefore the same spectra are analysed and according to formula (2.13) the product probabilities for the colour ratios and brightness temperatures are computed. The class assignment is effected by the highest resulting probability. Since we know the affiliation of the input spectra, a score value of correctness can be evaluated which then helps to select the optimal ratios by maximisation of the correct assigned spectra. The histograms of the optimal ratios for each class provide the basis for the classification method described in section 4.1.2.

2.4 Corrections and Plausibility Tests

In an early stage of the processing it is helpful to sort out suspicious or corrupted radiance profiles or to apply correction to parameters, which are essential for the later processing and validations steps. Some first important processing steps are described in the following sections.

2.4.1 Altitude correction

2.4.1.1 Correction of tangent altitudes

The engineering tangent altitudes connected with the L1B MIPAS data are known to have uncertainties up to several kilometers [19]. These errors are typically varying by about 1.5 km within one orbit (in case of ESA processor version $\geq$IPF/4.61 and $<$IPF/4.67). Thus, any cloud top determination algorithm based only on level 1b dataset is exposed to the same errors.

In order not to have to completely rely on ENVISAT space craft position and attitude and MIPAS scan mirror position information, methods have been developed to retrieve tangent altitude information directly from the spectra. The absolute pointing information can be retrieved either in terms of pressure at the tangent point [21], or in terms of geometrical tangent altitudes [20]. Both these methods retrieve the relative pointing information in geometric co-ordinates. The different principles behind these approaches, i.e. the tangent pressure retrieval and the retrieval of geometric tangent altitudes are both suitable to correct cloud altitude information.

In the following, we propose the implementation of three different methods to handle tangent altitude information in the processor.

2.4.1.2 No correction of engineering tangent altitudes

In this mode the engineering tangent height information is used without any correction. This mode is foreseen to handle level 1b data versions for which no external input as needed for the following two correction methods is available. These might be new level 1b data versions for which no retrieval results of tangent pressure from ESA level 2 processing nor results of geometric tangent altitude retrieval from the IMK processor are existing. This mode can also be suitable for different kind of validation purposes.

2.4.2 Correction by use of geometric tangent altitude retrieval

This method is based on pre-existing retrievals of geometric tangent altitude as available from IMK [20]. From these retrievals a database of mean tangent altitude corrections $\delta h_{iday,jlat} = (h^{fit} - h^{eng})_{iday,jlat}$ is calculated. These values are mean corrections over single days $iday$ for various latitude bins $jlat$. Additionally, $\delta h_{iday,jlat}$ are mean corrections over altitude, i.e. one number which is valid for all tangent altitudes.

This database serves as input quantity on basis of which the cloud processor performs a tangent altitude correction as follows:

$$h_{ih}^{corr}(t,x) = h_{ih}^{eng}(t,x) + \delta h(t,x) \tag{2.14}$$

where x is the actual latitude and t the actual time of the limb-scan. ih denotes the different sweeps of the limb-scan.

The value of $\delta h(t,x)$ is calculated by bi-linear interpolation based on $\delta h_{i,j}$, $\delta h_{i,j+1}$, $\delta h_{i+1,j}$, and $\delta h_{i+1,j+1}$ where $x_i \leq x < x_{i+1}$ and $t_j \leq t < t_{j+1}$.

2.4.2.1 Correction by use of pressure at tangent point retrieval

Here we use the information from the ESA MIPAS level 2 processing of pressure at tangent points which is available for each limb-scan, and, thus, does not require any interpolation as the previous method. Thus it is considered to be the most accurate scheme. As input quantities this method needs (a) level 2 results of tangent pressure and (b) geopotential height H and pressure P from the ECMWF analysis.

The following scheme is proposed:

- Extract from the ECMWF analysis H^{ecm} and P^{ecm} for the location and time of the actual limb-scan by interpolation in latitude, longitude and time.
- Determine the geopotential heights H_{ih} of each sweep ih of the limb-scan for which pressure has been retrieved and is available in the level-2 dataset by interpolation in $\ln P$ space:

  $$H_{ih} = H_i^{ecm} + \frac{\ln P_{ih} - \ln P_i^{ecm}}{\ln P_{i+1}^{ecm} - \ln P_i^{ecm}} \left(H_{i+1}^{ecm} - H_i^{ecm}\right) \tag{2.15}$$

  with $P_{i+1}^{ecm} \geq P_{ih} > P_i^{ecm}$.
- Determine the geometric tangent altitudes by application of the formula:

  $$Z(H,\Phi) = [1 + 0.002644\cos(2\Phi)]H + [1 + 0.0089\cos(2\Phi)]\frac{H^2}{6245\,km} \tag{2.16}$$

  (M.J. Mahoney, http://mtp.jpl.nasa.gov/notes/altitude/altitude.html)

  where Z is the geometric altitude, H the geopotential height, and Φ the latitude.
- Determine the geometric tangent altitudes of the sweeps for which (e.g. due to cloud contamination) no pressure retrieval has been performed by using the differences between the engineering tangent altitudes Z^{eng}:

  $$Z_{ih+1} = Z_{ih} + Z_{ih+1}^{eng} - Z_{ih}^{eng} \tag{2.17}$$

The priority of correction is that no altitude correction is the default. If L2 and ECMWF/climatological pressure data is present, altitude is corrected by use of pressure at tangent point retrieval. If no L2 data is present but time is in a certain range the Kiefer et al LOS correction is used.

2.4.3 Colour Ratio Tests

The cloud index (CI) value is a frequently used tool in the MIPAS cloud retrieval. It is a simple radiance ratio for two micro windows, also more generally called colour ratio (CR). Amongst other things colour ratios are able to identify and assign suspicious spectra in the input L1B data. Detailed analyses with the MIPAS standard cloud index for band A (CI-A), with the MWs

$$MW_1 := [788.20 - 796.25]\, cm^{-1}$$
$$MW_2 := [832.30 - 834.40]\, cm^{-1},$$

offered the typical probability distribution function (PDF) (Figure 2.1). The physical background for this distribution with altitude is very well known (see also section 3.3). Therefore CI-A is best suited for a simple plausibility test by the following steps:

(1) One has to calculate the CI profile of the MIPAS scan, by the quotient of the two radiance profiles for the pre-defined two MWs.

$$CI_{prof} = prof_1(MW_1)/prof_2(MW_2). \tag{2.18}$$

To avoid numerical overflow the profile radiances have to be checked to be greater than a given minimum value (rad_{min}). The minimum value chosen in this application is $10^{-19}\ W/cm^2\, sr\, cm{-}1$. Values smaller rad_{min} need to be converted to rad_{min}(see also section 3.1).

(2) Checking if CI(i) below/above certain threshold values, which are defined as $CI_{min} = 0.5$ and $CI_{max} = 50$. Single CI values can be significantly above or below the thresholds, when the radiance values are close to the noise level (e.g. 50 $nW/cm^2\, sr\, cm^{-1}$). However below 30 km altitude CI-A values below 0.5 are suspicious. If more than one sweep below 30 km show $CI < CI_{min}$ the complete profile should be rejected from further analysis.
The L1B sweeps where the restriction is violated are flagged as “not valid” and are rejected from further analysis.
The flag value differs from the MIPAS L1B quality flag “corrupted” so that they are distinguishable.

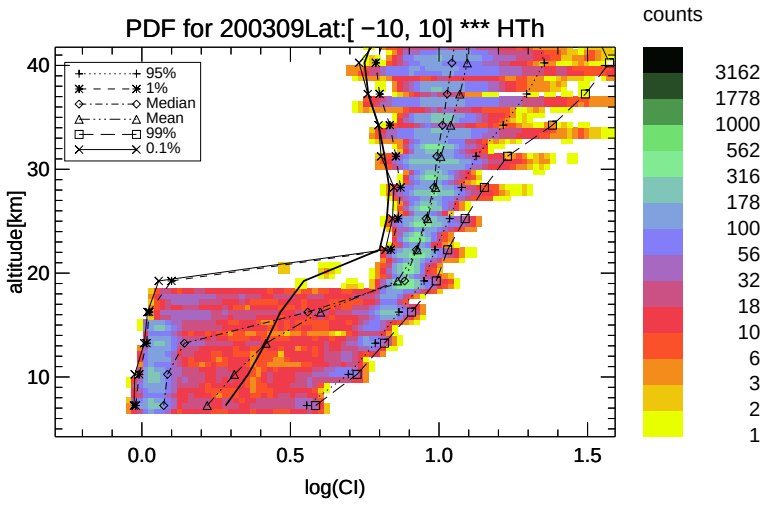

Figure 2.1: PDF distribution for CI-A for September 2003 in the equator region.

2.4.4 Temperature Criteria

Not only for validation of cloud measurements it is good to know the maximum cloud formation temperature in the altitude range of interest. For example, if the temperatures derived from ECMWF at the observing location are well above this threshold temperature, clouds are very unlikely to occur. As a basis for this estimated threshold the average, minimum, and maximum profiles for the equatorial region, mid-latitudes, polar summer and winter of the reference atmospheres databases by J.J. Remedios are used. The cloud formation temperatures for three different cloud types are calculated. For the liquid water cloud formation temperature the formula of Bolton (Bolton, D., The computation of equivalent potential temperature, Monthly Weather Review, 108, 1046-1053, 1980) is used. These is valid for temperatures between 243 - 308 K. The ice cloud formation temperature is calculated according to Mauersberger and Krankowsky (GRL, 10.1029/2002GL016183, 2003), valid for 273.15 K > T > 164.5 K. The equation for the NAT cloud formation temperature is from Hanson and Mauersberger (GRL, vol 15, no8, pp855-858, 1988).

As an example the threshold temperatures for the equatorial region are shown in Figure 2.2. The solid black, blue and red lines are the temperature profiles for the average, minimum and maximum case respectively. The dotted lines denote the liquid, the dashed lines the ice and the dash-dotted lines the NAT cloud formation temperature thresholds.

On the right hand side of Figure 2.2 the maximum threshold temperatures for the equatorial region (black), mid-latitudes (red), polar summer (blue) and winter (green) are presented. The solid grey line highlights the all region maximum envelope ($T_{max,cloud}$). This curve can be used for a plausibility temperature test. Spectra detected as cloudy a few K above the threshold will be flagged as *warm cloud* and spectra well above the threshold ($\Delta T_{thres} > 5K$) should be marked as *suspicious cloud* or just rejected from the follow-on analysis.

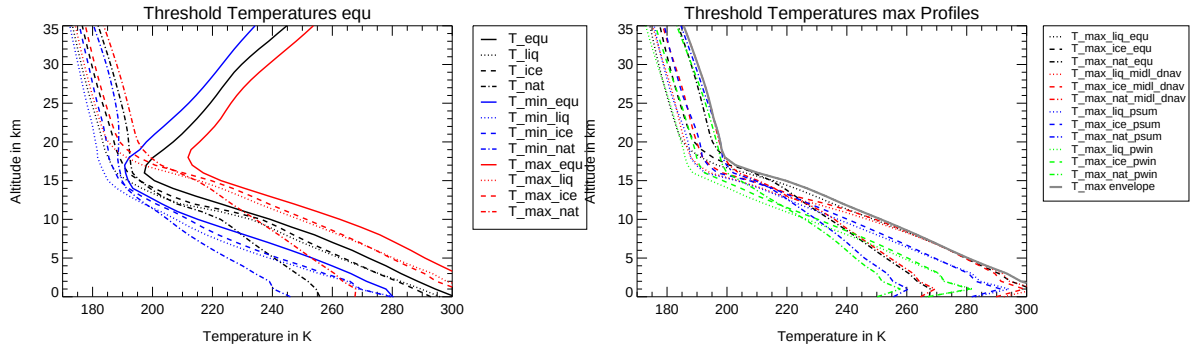

Figure 2.2: Threshold temperatures for ice, NAT and liquid formation with altitude for various conditions: equator (left) for minimum and maximum trace gas contribution and 4 reference atmospheres (right) only maximum conditions with superimposed envelope profile in grey.

3 Cloud Detection

3.1 Operational Cloud Index Method

For the detection of clouds in the MIPAS L1B thermal emission spectra, a fast and effective cloud detection method has been required through the mission to identify cloud-contaminated MIPAS sweeps. Historically, this function has been fulfilled by the Cloud Index (CI) method based on the method originally described in [8] and extended for MIPAS in [10]. Although its original purpose was to remove cloudy spectra from trace gas retrievals, the CI method has also been used successfully to derive cloud distributions/occurrence frequencies from MIPAS (e.g. [14], particularly in band A).

The CI method is founded on the simple relation for a colour ratio:

$$CI = \frac{\frac{1}{n_1}\sum_{i=1}^{n_1} rad_i(MW_1)}{\frac{1}{n_2}\sum_{i=1}^{n_2} rad_i(MW_2)} \tag{3.1}$$

with rad_i the radiance measured at the wavenumber increment i of the corresponding MW.

Operational CI cloud detection, as implemented for MIPAS either at ESA or at off-line processing groups, utilises this ratio of the integrated radiances in two microwindows (where microwindow 1 is 'MW1' and is a control microwindow, microwindow 2 is 'MW2' and so forth). For example, for MIPAS band A, the ratio is calculated as the mean radiance in the spectral range of 788 - 796 cm^{-1} (MW1) over the mean radiance in 832 cm^{-1} - 834 cm^{-1} (MW2). This ratio is known as CI-A as it is specific to MIPAS band A.

Typically, values for CI-A are close to unity when an optically thick cloud is present in the MIPAS FOV and CI-A tends to be large (CI-A $>$ 6.0) for a clear sky line of sight. Low to high CI-A values represents the transition from optically thick to optically thin clouds present and therefore cloud information can potentially be retrieved from any sweep that possesses a CI-A from 1 to 6.

Other intervals in MIPAS bands B and D, CI-B and CI-D respectively, have been calculated and are shown also in Table 3.1. These CI microwindow pairs should also be implemented but CI-D should be treated with considerable caution.

These microwindows have been selected, given the tropical upper troposphere as a constraint and global testing. It is entirely reasonable to define other CI pairs which would provide cloud detection in particular regions, e.g. the polar lower stratosphere where the gas contribution is less complex and intense.

Cloud Index MIPAS Band	MW1 (cm^{-1})	MW2 (cm^{-1})
CI-A	788.20 - 796.25	832.3 - 834.4
CI-A2	788.20 - 796.25	926.0 - 932.0*
CI-A3	788.20 - 796.25	947.0 - 952.0*
CI-B	1246.3 - 1249.1	1232.3 - 1234.4
CI-D	1929.0 - 1935.0	1973.0 - 1983.0

Table 3.1: Cloud index MW pairs for MIPAS; *further optimisation possible

3.1.1 Complexity of thresholds

For all CI MW pairs, thresholds need to be defined.

1. In the CSDB, CI-A values are observed across the range from 1 to greater than 10.
2. Hence CI MW thresholds depend on trace gas variability.
3. CI thresholds are known to be altitude dependent in principle because of the trace gas dependence with altitude, particularly the sharp transition from troposphere to stratosphere.
4. CI thresholds should be latitude dependent for the same reason although this may be less important

Hence CI MW thresholds should allow for altitude and probably latitude dependence.

3.1.2 Masked microwindows

In theory the CI approach can also be improved by using masking of frequencies within a cloud detection window; this is performed operationally for MIPAS in the trace gas retrievals but not in the cloud detection approach:

1. Masking of microwindows is performed to remove the trace gas signatures from the chosen microwindows so that the radiance in that microwindow from the underlying cloud is maximised.
2. Any trace gas variability that may cause uncertainty in the microwindow and furthermore affect the efficiency of the microwindow for use in cloud detection can be optimised to be as low as possible.
3. The masked microwindow is calculated off-line with an expert approach (see following sub-section):

$$MW_{masked} = MEAN(MW(Mask)) \tag{3.2}$$

$$MW_{unmasked} = MEAN(MW) \tag{3.3}$$

3.1.3 Microwindow variance

1. Statistical approach using real MIPAS level 1b data
2. Split chosen microwindow into small increments (n)
3. Calculate the mean radiance over each increment

4. Calculate the standard deviation of all the data in each increment
5. Ratio of standard deviation/mean radiance (sdev/mrad) shows regions where variability is high due to trace gas variation.

3.2 Optimised CI Method

The Optimised CI Method [15] basically consists of the CI Method described in Section 3.1 run with different microwindows, which have been optimised so that the CI best correlates with the amount of cloud present in the MIPAS FOV, as parameterised by the cloud effective fraction EF.

EF is defined as

$$\mathrm{EF} = \frac{\int_{-d}^{z_{\mathrm{ct}}} (1 - e^{-k_{\mathrm{ext}} x}) \phi(z) dz}{\int_{-d}^{d} \phi(z) dz} \tag{3.4}$$

for a FOV of width $2d$ characterised by the convolution function $\phi(z)$ corresponding to integrated pencil beam radiances each penetrating a pathlength x through an atmosphere of extinction coefficient k_{ext} and cloud top height z_{ct} relative to the tangent height. It is essentially the effective blocking power of the cloud within the FOV — the proportion of the FOV filled by cloud multiplied by the extinction of the cloud. Therefore, an EF$= 0$ indicates that a measurement is cloud-free or clear, an EF$= 1$ represents a FOV that is completely filled with thick cloud and $0 <$EF< 1 represents the spread of varying cloud-filled states of a FOV.

The optimised microwindows (MW1 and MW2) for the Optimised CI Method are:

- MW1 $=$ [777.0,779.0] cm^{-1}
- MW2 $=$ [819.0, 820.0] cm^{-1}

and they represent an improvement in correlation of CI with EF of 13.8% in comparison with the original CI microwindows discussed in Section 3.1.

The Optimised CI test is applied to MIPAS data one scan at a time. The optimised CI is evaluated at each sweep in the scan pattern, from 30 km downward. If at any point in the scan pattern the calculated optimised CI exceeds the threshold (often taken as 1.8, as in the ESA operational CI method), cloud is flagged. The highest altitude at which this exceedance criterion is satisfied is loosely called the cloud top.

3.3 CI Threshold Profile Method

A climatology for 2003 of the probability densities for the standard MIPAS Band A cloud index is the basis for this method. An example for July 2003 is shown in Figure 3.1 with different superimposed percentiles.

Obviously a simple mean or median profile would not represent a best estimate of the CI-threshold profile. Therefore the procedure is optimised by applying different criteria above, at, and below the altitude step, where the PDF starts to become a bi-modal distribution (see also Fig. 2.1).

The database consists of altitude profiles of threshold values from 3 to 42 km on a 1 km grid and 20 deg latitude bins from -90 to $+90$ degrees for the months January to December. Linear interpolation in time and latitude space shall be possible.

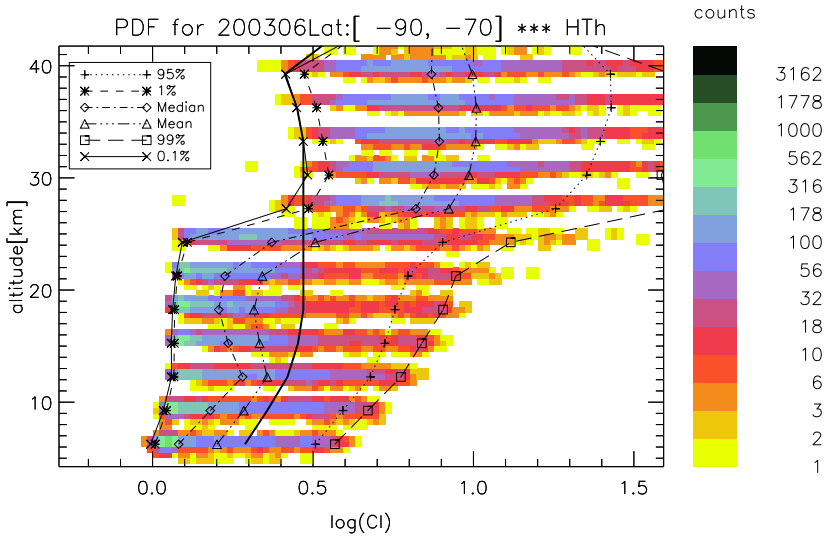

Figure 3.1: PDF analysis for CI-A versus altitude for June 2003 for -90° to -70° latitude. CI-Threshold profile is superimposed as a thick black line.

The analysis has compiled a climatology of threshold profiles. The latitudinal and temporal evolution of the profiles is shown in figure 3.2 for two examples (April and -90 to -70 deg latitude respectively). The threshold method is analysing a profile by moving top down through all tangent heights. A measured spectrum is flagged as cloudy if:

$$CI_A < CI_{thres}(alt, lat, month) \quad \Rightarrow \quad \textbf{Cloud Flag = FLAG}_C(alt) \; = \; 1$$

For that procedure it is necessary to interpolate CI_{thres} to the tangent height of interest by linear logarithmic (log CI) interpolation.

Similar procedures will be used to improve the CI_A, CI_B and CI_D thresholding. Therefore it is helpful to create a simple interface for processor, where MWs can be defined and CI-threshold values depending on latitutde, alitude and time can be read in.

3.4 Simple Radiance Thresholding Method

A very basic method is the Mean Radiance Threshold test which simply uses a statistically-gathered threshold upon radiance to detect cloud by assuming that clouds have a warmer brightness temperature than a clear limb view — and hence measurements containing cloud will have a higher average radiance than clear measurements. The detection test then consists of evaluating the average radiance $\bar{L}$ in a chosen microwindow in the MIPAS spectra and comparing it against a threshold Thr, such that:

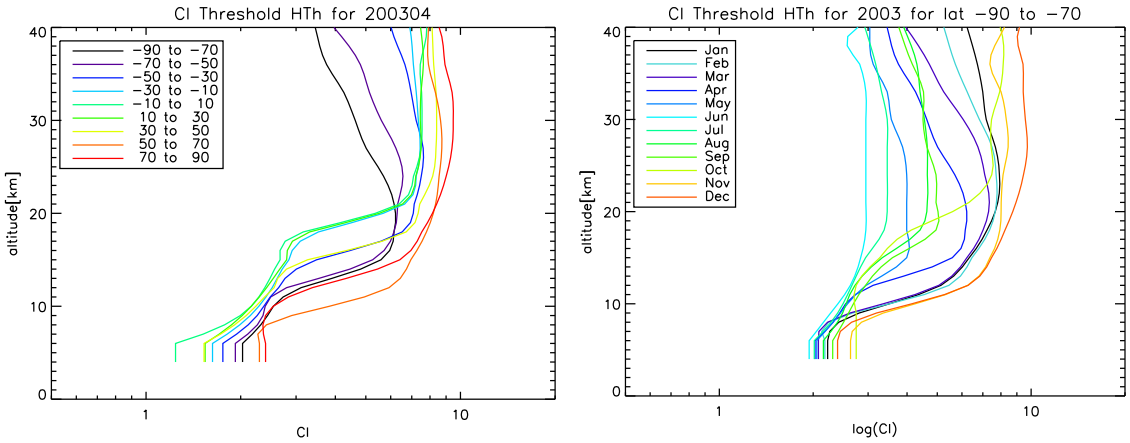

Figure 3.2: Latitudinal (left) and temporal (right) evolution of the retreived CI-A threshold profiles.

Altitude	*Threshold* $[nW/cm^2srcm^{-1}]$
6.0	758.565
9.0	172.940
12.0	170.634
15.0	93.8932
18.0	82.1588
21.0	86.9026
24.0	88.4796
27.0	89.1483
30.0	89.3419

Table 3.2: Altitude-dependent thresholds for the simple radiance thresholding method.

- if $\bar{L} < Thr$, then measurement is CLEAR
- if $\bar{L} \geq Thr$, then measurement is CLOUDY.

For MIPAS, one considers the region around 960.7 cm^{-1} (960.0 – 961.0 cm^{-1}) as it has high transmission and low gaseous emission. The threshold Thr is chosen using an ensemble of radiative transfer model simulations (the Reference Forward Model RFM, a GENLN2-based line-by-line originally developed by Dudhia at Atmospheric, Oceanic and Planetary Physics, Oxford University which provides reference spectral calculations for MIPAS [13]) — and is highly user-function dependent. For instance, if a user requires that a cloud having an extinction coefficient of 10^{-4} km^{-1} be detected, a threshold of 100 nW/cm^2 sr cm^{-1} must be chosen at a tangent height of 9 km (higher threshold for lower tangent heights and for higher extinction values).

The Simple Radiance Thresholding test is applied to MIPAS data one scan at a time. The average radiance in the given microwindow is evaluated at each sweep in the scan pattern, from 30 km downward. If at any point in the scan pattern the calculated average radiance exceeds the threshold, cloud is flagged. The highest altitude at which this exceedence criterion is satisfied is loosely called the cloud top.

The thresholds are altitude-dependent, and computed by considering the average radiance in the $960 - 961 cm^{-1}$ MW for 2003 MIPAS data as a function of height. They are then determined by considering probability density functions of the data at each tangent height. The values are collected in Table 3.2

3.5 Singular Vector Decomposition Method

SVD is a standard statistical technique [16] used for finding patterns in high dimensional data and for summarising this data. To this end, SVD transforms a number of potentially correlated variables into a smaller number of uncorrelated variables called singular vectors **SV**. The first singular vector $\mathbf{SV}_1$ accounts for as much of the variability in the data as possible, and then each successive $\mathbf{SV}_i$ accounts for as much of the remaining variability as possible.

An ensemble of RFM-simulated MIPAS spectra containing varying amounts of cloud have been used to define singular vectors which span the clear and cloudy atmospheric states [15], called $\mathbf{SV}_{\text{clear}}$ and $\mathbf{SV}_{\text{cloudy}}$, respectively. The simulated spectra — and hence the singular vectors cover the upper half of the MIPAS A band (a spectral range of 827.5 – 970 cm^{-1}) because the bottom half of the MIPAS A band is characterized by strong gas lines. These singular vectors have been defined for each tangent height in the MIPAS nominal scan pattern.

Any arbitrary spectrum can be successfully fit to a high degree using this set of altitude-dependent singular vectors which span the clear and cloudy atmospheric states. Using the previously calculated clear and cloudy singular vectors, $\mathbf{SV}_{\text{clear}_i}$ and $\mathbf{SV}_{\text{cloudy}_i}$ (of which there are m_{clear} and m_{cloudy}, respectively), for each MIPAS TH where cloud is normally expected, any measured MIPAS spectrum in the spectral range of 827.5 – 970.0 cm^{-1} can be accurately fitted by a linear least squares fit in the singular vectors. Taking an arbitrary MIPAS IR spectrum L_{orig}, the first step is to normalise the spectra by subtracting its average radiance ($\bar{L}_{\text{orig}}$) such that

$$L_{\text{norm}} = L_{\text{orig}} - \bar{L}_{\text{orig}}. \tag{3.5}$$

The linear least squares fit L_{fit} of this normalised spectrum L_{norm} is then trivially found, such that

$$L_{\text{fit}} = \sum_{i=0}^{m_{\text{clear}}} \lambda_{\text{clear}_i} \mathbf{SV}_{\text{clear}_i} + \sum_{i=0}^{m_{\text{cloudy}}} \lambda_{\text{cloudy}_i} \mathbf{SV}_{\text{cloudy}_i}, \tag{3.6}$$

where λ_{clear_i} and $\lambda_{\text{cloudy}_i}$ are constant coefficients of the least squares fit.

Once this linear least squares fit has been obtained, it is trivial to reconstruct the radiance components of the original signal: that due to the clear background state and that due to possible cloud presence. Reconstructing, the clear radiance is

$$L_{\text{clear}} = \sum_{i=0}^{m_{\text{clear}}} \lambda_{\text{clear}_i} \mathbf{SV}_{\text{clear}_i}, \tag{3.7}$$

and the radiance due to the cloud presence is

$$L_{\text{cloudy}} = \sum_{i=0}^{m_{\text{cloudy}}} \lambda_{\text{cloudy}_i} \mathbf{SV}_{\text{cloudy}_i}. \tag{3.8}$$

It follows, then, that when the radiance due to cloud presence becomes non-zero, cloud is present. To normalise this quantity, the ratio of the cloudy radiance to the total radiance L_{total}, called the Integrated Radiance Ratio, is considered such that when

$$\frac{\bar{L}_{\text{cloudy}}}{L_{\text{total}}} > 0 \tag{3.9}$$

for cloudy spectra and

$$\frac{\bar{L}_{\text{cloudy}}}{L_{\text{total}}} \approx 0 \tag{3.10}$$

Altitude	*Threshold*
7.5	-0.769068
8.5	-0.834624
9.5	-0.817742
10.5	-0.907642
11.5	-1.01306
12.5	-1.06507
13.5	-1.07146
14.5	-1.00277
15.5	-0.913329
16.5	-0.902765
17.5	-0.938599
18.5	-0.938599
19.5	-0.938599

Table 3.3: Altitude-dependent thresholds for SVD Integrated Radiance Ratio method.

for clear spectra, where $\bar{L}$ represents the average of the reconstructed radiance in the 960 cm^{-1} – 961 cm^{-1} MW. The logarithm of the Integrated Radiance Ratio is the metric which is thresholded in this method.

A set of altitude-dependent thresholds have been statistically defined from application of this test to real full-resolution MIPAS data, as shown in Table 3.3.

In the same manner as previously discussed for the other detection methods, this method is applied one scan at a time. The Integrated Radiance Ratio is evaluated at each sweep, using the appropriate set of altitude-dependent singular vectors in the fitting process, and once the logarithm of the Integrated Radiance Ratio exceeds the altitude-dependent threshold, then cloud is said to occur.
If altitudes $z_{l1b} \leq z_{max_{thresh}} + \frac{\Delta z}{2}$ with $\Delta z = z_1 - z_0$ the method is applied where the threshold next to the tangent point altitude is chosen (If $z_{l1b} < z_0$ also t_0 is chosen).

3.6 Water Vapour Contiunum Proxy

Analyses in [1] showed that simple CI col or ratio CI-A is quite sensitive to tropospheric water vapour if the mixing ratios are greater than 1000 ppmv. It is enhancing the continuum and is very difficult to differentiate from the cloud continuum. In consequence, high water vapour makes the detection of clouds with CI-CR ambiguous at altitudes with water vapour above 1000-3000 ppmv. Two information would help to solve this dilemma.

1. An information based on the MIPAS spectra about a rough estimate of the water vapour mount (e.g. by a CR or BTD proxy.
2. An lower altitude threshold where the water continuum starts to get problematic in sense of cloud detection

3.6.1 Water Vapour Index

Analyses in [1] and references therein showed that a BTD of one MW of water vapour line and a close-by MW for the background emission is an excellent proxy for tropospheric water vapour, with:

$$\begin{aligned} MW_1 &:= [784.0 - 784.8]\, cm^{-1} \\ MW_2 &:= [787.0 - 788.0]\, cm^{-1}, \end{aligned}$$

The Water Index (WI) is well positive correlated with water vapour for each background atmosphere (latitudes). But for extreme high values (>1000 ppmv) BTD is shrinking, which makes the differentiation between high water vapour continuum and clouds difficult again. This takes place for altitudes $>\sim$ 6 km below the tropopause. A simple flagging like:

$$WI > WI_{thres}(alt, lat) \quad \Rightarrow \quad \textbf{FLAG}_{WI} = (-)1$$

will help to sort out doubtfull cloud cases in scientific analyses.

3.6.2 Altitude Restriction for Cloud Detection

The question at which altitudes one can expect such high values water vapours values can not be predicted with high accuracy due to the varying atmospheric conditions and strong variability for water vapour in the troposphere. However, a restriction in cloud detection related to the tropopause height should be taken into account in the processing by simply flagging altitudes below this threshold altitude $h_{CD,min}(lat)$ as a critical altitude for cloud detection.

$$TH_i < h_{CD,min}(lat) \quad \Rightarrow \quad \textbf{FLAG}_{CDmin} = (-)1$$

Corresponding threshold values will be provided by the control file.

3.7 Scatter Effect Index

The Cloud Scattering Index (CSI) based on reference [9] will be applied to the MIPAS spectra for information on scattering effects in the measured spectrum. These effect is more pronounced for larger particles. Using MWs in the region of an absorbing H_2O line and a region close to it, then BTD(MW_1, MW_2) is negative if an absorption features occurs in the spectrum and detection of scattering is possible.

Following wavenumber regions have been selected for CSI^a:

$$\begin{aligned} MW_1^a &:= [803.50 - 803.60]\, cm^{-1} \\ MW_2^a &:= [803.70 - 803.90]\, cm^{-1}, \end{aligned}$$

and a second MW-pair MW^b with slightly more sensitivity in respect to S/N in the 950 cm^{-1} region as CSI^b::

$$MW_1^b := [948.225 - 948.325]\, cm^{-1}$$

$$MW_2^b \quad := [948.600 - 949.000]\, cm^{-1},$$

If BTD of the MW pairs MW^a and MW^b are negative for a detected cloudy sweep then flagging of the spectrum take place.

$$CSI^i = BTD(MW_1^i, MW_2^i) < CSI_{thres}^i \quad \Rightarrow \quad \mathbf{FLAG}_{CSI^i} = 1$$

The value of CSI_{thres}^i will be close to zero. But the exact value depends on the S/N of the detector band and will be specified in the control file. The size of the BTD of CSI includes information of particle radius and is stored in the processor expert-output file.

3.8 Decision

3.8.1 Confidence Criterion

Every presented cloud detection method has its uncertainties. A cloud flag is set for every analyzed sweep and each methods cloud index. Now a combination of this results is measured to provide a global cloud decision and confidence. For all methods and cloud indices a certain weight w shall be defined in the CPF. The detection result flags are combined with the weights and the cloud confidence for a certain tangent height and can be written as a weighted sum over all detection results (detection methods and each having a different number of involved CIs)

$$\mathrm{CONF}_{cloud} = \sum_{i=0}^{n_{CD}} \mathbf{FLAG}_{CD_i} \cdot w_{CD_i} \tag{3.11}$$

The normalised confidence as a global cloud indicator is given by

$$\mathrm{NCONF}_{cloud} = \frac{\mathrm{CONF}_{cloud}}{w_{sum}} \in [0,1]. \tag{3.12}$$

where w_{sum} is the sum of all weights of the detection methods and their CIs

$$w_{sum} = \sum_{i=0}^{n_{CD}} w_{CD_i}. \tag{3.13}$$

This confidence is used to identify a cloudy sweep iff NCONF_{cloud} is larger than a defined threshold T_{nconf}. In addition cloud confidence levels/ classes will be specified by deviding $[0,1]$ into subsets. Table 3.4 shows an example of this class definitions as they are set in the actual control paramater file.
The current value for T_{nconf} is 0.2 which is also set as the transition value from “disputable” to “likely” cloud. This setting is not necessary but reasonable.

Moreover the cloud top longitude and latitude is determined for every scan. If a scan is analyzed to be “cloud free” the longitude and latitude are the mean values over the scan.

3.8.2 Macro Retrieval Results

If the Macroscopic Parameter Retrieval measures a cloud top height value this one is used for the detection summary as it is meant to be a more precise result. The detection summary is then modified by this result.

normalized cloud confidence value	*cloud confidence class*
$\{0\}$	clear sky / cloud free
$]0, 0.2[$	disputable cloud
$[0.2, 0.5[$	likely cloud
$[0.5, 0.8[$	very likely cloud
$[0.8, 1]$	confident cloud

Table 3.4: Normalized Cloud Confidence Classes

The valid altitude neighbours of the retrieved cloud top height are used for interpolation of cloud top latitude, longitude and macro retrieval pressure. If only one neighbor sweep is available, values are simply transferred.

If the retrieved cloud top altitude is greater than the highest “cloudy” altitude on the L1B grid and the nearest neighbour altitude is flagged as “cloud free” the cloud flag is modified to “cloudy” and the corresponding normalized cloud confidence for this sweep is set to the confidence threshold value T_{nconf}. If sweeps above the related L1B altitude of the measured cloud top are flagged as “cloudy” these are modified by changing the cloud flag to “cloud free” and setting the normalized confidence to the smallest defined value, e.g. 0 with confidence class “clear sky”.

4 Classification

4.1 Introduction

A classification scheme of the cloud type which is based only on the IR spectra would significantly reduce the complexity to retrieve microphysical parameter. The classification of the spectra in respect to the most likely composition of the events is already a valuable product for scientific analyses. For example the occurrence for PSC types in the MIPAS observations over various winter gave already new insights in the formation mechanisms of Arctic and Antarctic PSCs (e.g.[5], [6] and [11], [12]).

4.1.1 BTD and Colour Ratio Approach

The BTD, colour ratio (CR) and combined BT-CR approach are based on a very simple technique. Due to different gradients in the continuum like cloud emissions the radiance ratios of MW pairs or the differences in the equivalent BT of the MW-emissions can behave very characteristically for special cloud types.

A correlation or scatter diagram of two parameter (like CR_1 vs CR_1 or BTD_1 vs BTD_2) can show clear separation in different region of the parameter space for different cloud types, which allows a distinctive classification in $Type_1$ to $Type_N$. However, frequently this classification is not unambiguous for all classes of possible cloud spectra or ambiguous classification is only possible for spectra under specific conditions (e.g. optically thick/thin conditions or spectra in the cloud top region). Separation lines in the two parameter space can define the membership for a parameter pair of the measured spectrum in a specific cloud class. Results on the classification are given in the following sections.

General restriction to the methods are:

1) $CI_A > 6$ spectra are extreme events in CSDB model runs with low volume densities or integrated area density path (ADP, see also section 6.2). These spectra can not be differentiated from cloud free spectra with contribution from a realistic background aerosol and should be no considered in the cloud classification.

2) Cloud classification should be only performed if the CI-A values are in a range of a typical CI-A probability distribution (e.g. $0.5 < CI_A < 6$).

3) The classification needs to be cross-checked with the temperature at the tangent point, if T is realistic in respect to the cloud formation processes.

Corresponding threshold values for CI-A and T are defined in section 2.4

4.1.2 Primitive Bayes Classifier Method

Basis for the analysis are the histograms of the cloud classes explained in section 2.3.3 and the selected colour ratios and BTD of the Bayes classifier. Figure 4.1 show the histograms for two ratios on a brightness temperature based analysis for NAT/ice and NAT/STS classification by the simulated spectra from the CSDB. Large overlap in the histograms indicates more difficulties in classification.

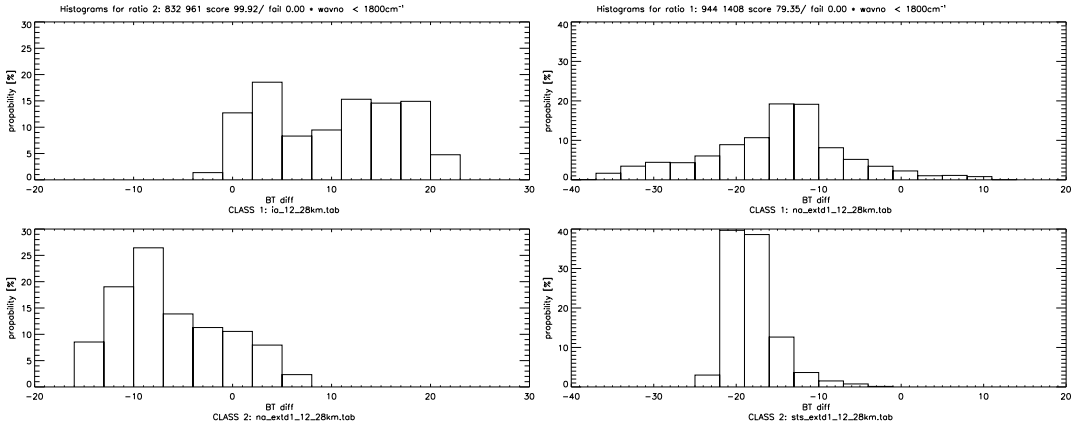

Figure 4.1: Bayes histograms for ice/NAT (left) and NAT/STS (right) classification.

Each brightness temperature difference of an input spectrum can be attributed to a specific probability p_i in each class (cloud type) All selected brightness temperature differences then build up a product probability

$$P_j = \prod_{i=1}^{N} p_{i,j}$$

for each potential cloud class (j: histogram class) over all selected MW-pairs (i: number of histograms). The maximum in P_j explains the affiliation to the cloud type class j. As a consequence it is possible to assign every single spectrum to a certain input cloud type.

The maximum probability $P_{j,max}$ does not have any significance to the validity of the classification. Hence, a statement on the quality of naive Bayes cloud classification has to be given separately. One idea is to involve the separation of the histograms. If the histograms show a distinct separation, i.e. they cover different brightness temperatures, the result would be a hundred per cent assurance of classification. Otherwise the accumulated overlapping histogram area gives information on the “reliability”. As a rough rule of thumb, the more overlapping the greater the uncertainty of cloud classification.

4.2 Stratospheric Clouds

4.2.1 NAT identification

The NAT detection follows the analyses by [9] and [5]. Two colour ratios, the operational CI-A and a so-called NAT index (NI) can be separated in respect to NAT particles with small radii ($R_{mean} < 3\mu m$ as illustrated in

4.2.

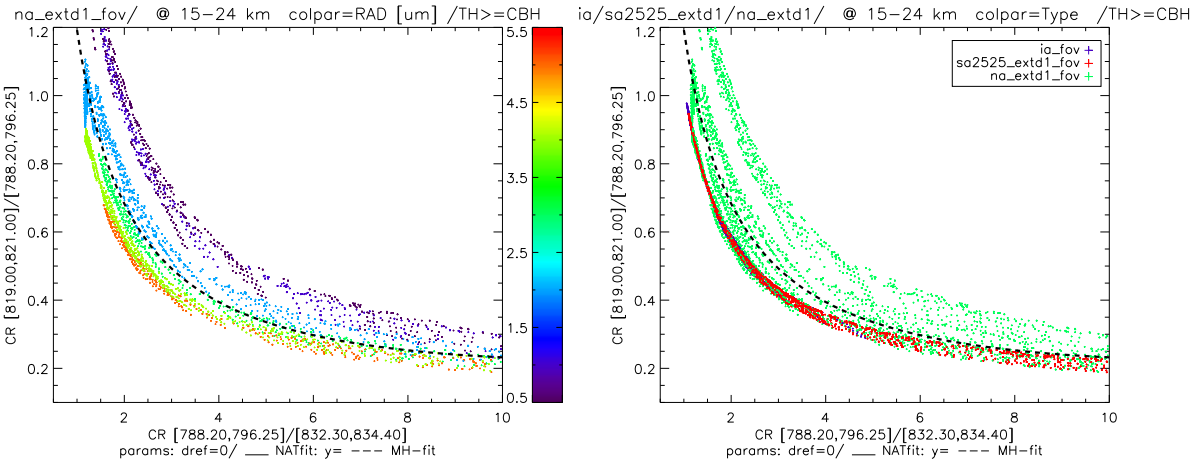

Figure 4.2: Correlation of CI-A versus NAT index for (left) CSDB spectra of NAT scenarios and (right) scenarios for NAT, STS and ice particles in the altitude range 15-25 km. The MIPAS FOV and only spectra where $TH > CBH$ were considered.

A NI-threshold function (NI_{thres}) has been fitted to the data. The function is valid over a broad altitude range. Some improvements to the fit by some adjustment at the top and bottom level (12-15 km and >25 km) might help for an even more sensitive separation, but are not absolutely essential.

$$NI_{thres}(CI_A) = (0.1536 + 0.71531 \cdot CI_A - 0.03003 \cdot CI_A^2)^{-1} \tag{4.1}$$

with $CI_A = rad(788.20 - 796.25\,cm^{-1})/rad(832.3 - 834.4\,cm^{-1})$

The curve can be applied in the CI-A range from 0.5 to 6, which covers optically thick to thin conditions. $CI_A > 6$ spectra are extreme events with low cloud amount in CSDB model runs and can not be differentiated from cloud-free spectra (see above). For classification one needs to compute CI_A and NI for the measured spectrum and:

$$NI > NI_{thres}(CI) \quad \Rightarrow \quad \textbf{FLAG}_{NI} = 1$$

4.2.2 Ice/STS and Ice/NAT differentiation

Similar to the NI approach it is possible to classify between Ice and STS even more effectively than for NAT. Figure 4.3 correlates CI-A with the BTD for $MW1 = [832.3, 834.4]$ and $MW2 = [947.5, 950.5]$ for STS with ice and NAT with ice of all CSDB spectra. Obviously for STS an excellent separation takes place. A simple separation line ($ICESTS_{thres}$) is able to classify ice spectra with 100% certainty, despite a little number of ice spectra have an overlap with STS. For NAT the overlap regions becomes larger and in consequence the differentiation is more difficult. This is illustrated with the fit curve $ICENAT_{thres}$. Various BTD have been tested and the best pairs are presented here. From the third plot of Figure 4.3 it is obvious that STS and NAT behaves very similar for larger NAT particles ($R_{eff} > 3\mu m$) and small particles are responsible for the large overlap of NAT with with ice. Events (spectra) in between the STS-ice curve and the NAT-ice separation curve can be classified as

CI_A	*Threshold*
0.8	11.0
1.2	7.0
1.55	4.0
2.0	1.5
3.0	-1.5
4.0	-3.2
5.0	-4.7
6.0	-5.8
8.0	-7.7

Table 4.1: CI_A -dependent thresholds for Ice/NAT differentiation.

CI_A	*Threshold*
0.9	1.0
1.0	-0.5
1.5	-5.0
2.0	-7.0
3.0	-8.5
4.0	-9.5
6.0	-11

Table 4.2: CI_A -dependent thresholds for Ice/STS differentiation.

NAT for $NI > NI_{thres}(CI_A)$ and as ice for $NI < NI_{thres}(CI_A)$ (see section 4.2.1). In conclusion an Ice classification can be defined as:

$$BTD_{ICE}(833-948) > ICENAT_{thres}(CI_A) \quad \Rightarrow \quad \textbf{FLAG}_{PSCICE} = 1$$

or

$$ICENAT_{thres}(CI_A) > BTD_{ICE} > ICESTS_{thres}(CI_A) \ \textbf{AND} \ \textbf{FLAG}_{NI} = 0 \quad \Rightarrow \quad \textbf{FLAG}_{PSCICE} = 1$$

for $1 < CI_A < 6$ but over the complete altitude range of potential PSC formation.

Obviously NAT and STS are difficult to differentiate, especially for large NAT particles. Simple two parameter scatter schemes had only very limited success when testing various MW combinations. A set of CI_A- dependent thresholds have been statistically defined (see table 4.1 and 4.2) and the functions $ICENAT_{thres}(CI_A)$ and $ICESTS_{thres}(CI_A)$ are given by linear interpolation between the values.

For better classification between the three PSC types the naive Bayes classifier has to be applied.

4.2.3 Improved PSC classification

Improvements in the classification by BTDs can be achieved by using the simple Bayes classifier outlined in section 4.1.2. Main task of the Bayes analysis is the compilation of the histogram database with the best result on classification. For this an optimisation of the bin size needs to be applied, here the step size ΔBTD. Test have been made for $\Delta BTD = 1$ to 4 K.

To quantify the quality of the Bayes classification the CSDB spectra have been fed again into the Bayes classification program, separately for each cloud type, like described in section 4.1.2. For example, the

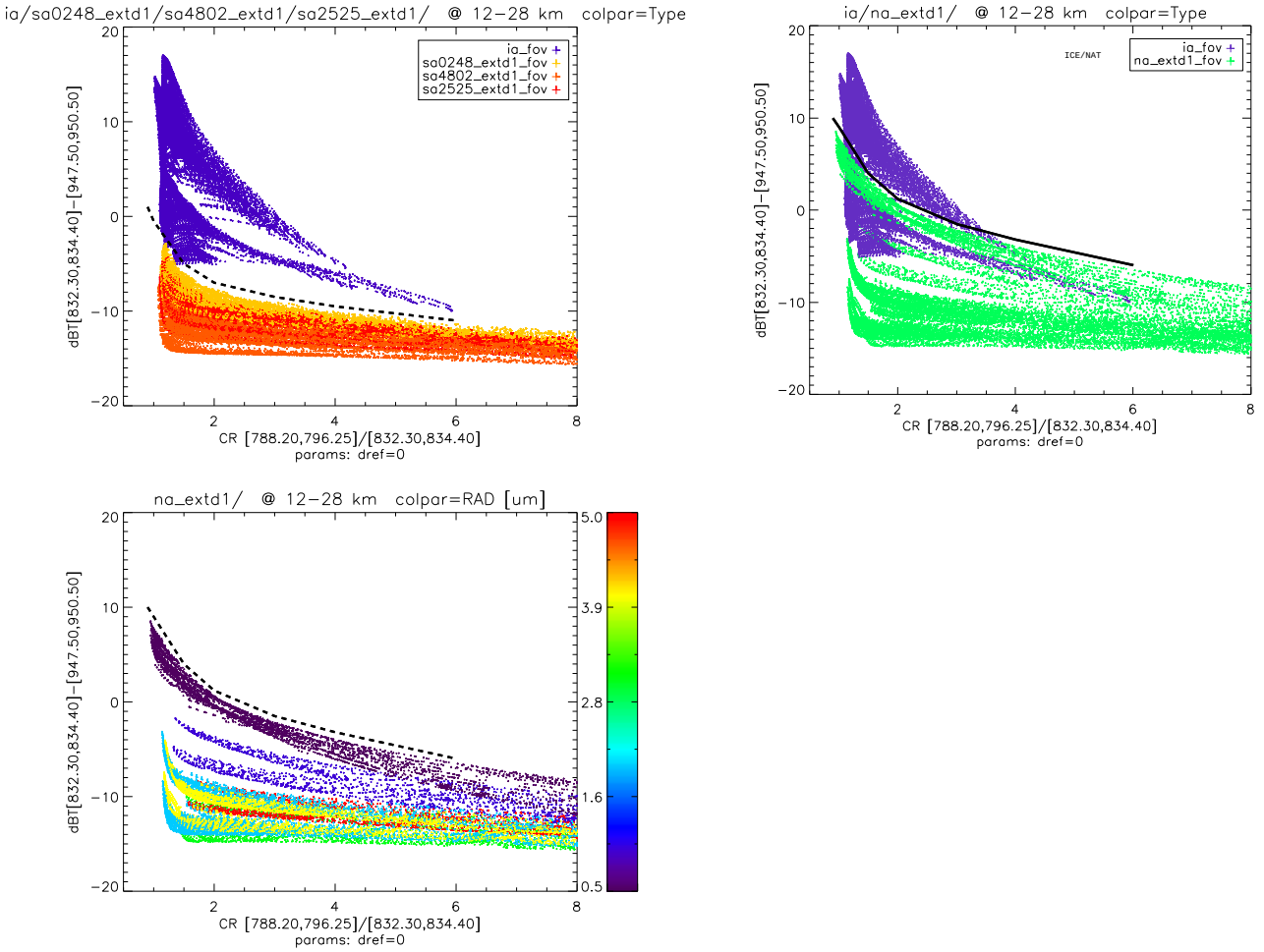

Figure 4.3: Cloud index versus BTD for CSDB spectra for (left) ice and STS, (right) ice and NAT, and (bottom left) only NAT with colour coded radius information

input of the ice clouds database should result in a 100% classification for ice. Smaller percentages give a rough marker for the quality or validity of the classification. If ice spectra are classified as NAT or STS, large percentages will highlight here problems for a unambiguous classification. Additional restriction to a sub-set of CSDB spectra in the histogram compilation and analysis, e.g. for optically thick/thin conditions by thresholds in CI-A or only spectra with tangent heights above the cloud base height (CBH) can help to improve the results. Various test have been applied.

An example is given in Table 4.3 for all database spectra and selected optically thick spectra. The classification works very well, but the restriction to optically thick events can improve the classification for NAT and one STS composition spectra from around 90 up to 100% and respectively 96 to 100% for sa0248.

The optimised histograms will be supplied by look-up tables. The corresponding MW-pairs for the computation of BTD will be part of the MW-selection database for $MW1_i, MW2_i$ for $i = 1, n$ (n_{max}=10). The classification follows the simple scheme described above (section 4.1.2).

ice/STS/NAT	Ice	NAT	sa0248	sa2525	sa4802
all	ia: 99.61 na: 0.39 sts: 0.00	ia: 0.00 na: 90.01 sts: 9.99	ia: 0.31 na: 3.23 sts: 96.46	ia: 0.58 na: 15.72 sts: 84.28	ia: 0.15 na: 52.28 sts: 47.72
thick	ia:100.00 na: 0.00 sts: 0.00	ia: 0.00 na: 100.00 sts: 0.00	ia: 0.00 na: 0.00 sts: 100.00	ia: 0.00 na: 0.00 sts: 83.03	ia: 0.00 na: 2.85 sts: 31.91

Table 4.3: Example of Bayes test analysis for ice-STS-NAT with all CSDB spectra and only for optically thick conditions ($CI_A < 1.2$). Various input test data sets classification in percent of the total number of spectra (ice: ia, NAT: na and various STS: sa compositions) for histograms with 4 K bins.

4.3 Tropospheric Cloud Types

The differentiation between liquid clouds and cirrus clouds or ice clouds in general is much more difficult then the PSC classification. Restriction to special conditions are necessary to get adequate test results, where a similar test approach like for the PSC classification were used.

4.3.1 Temperature criterium

In the upper troposphere and lower stratosphere temperatures are to cold for the formation of liquid clouds. Liquid clouds can only form for temperatures above -40°C. In consequence a retrieved cloud top temperature CTT $< -45°$C =$T_{\text{Liq, min}}$ corresponds to an ice cloud (taken into a count an uncertainty of 5 K in the retrieved CTT).

$$CTT < T_{Liq,min} \quad \Rightarrow \quad \textbf{FLAG}_{TnoLiq} = 1$$

and for each cloudy altitude step (sweep) with ECMWF temperature at the tangent point (TP):

$$T_{ECM}(TP) < T_{Liq,min} \quad \Rightarrow \quad \textbf{FLAG}_{TnoLiqECM} = 1$$

CTT significantly warmer than the extreme threshold temperature for cloud formation $T_{max,cloud}$ (section 2.4.4) have to be flagged as *warm clouds*, and should be handled with caution in further analysis.

$$T_{ECM}(TP) > T_{max,Cloud} + \Delta T \quad \Rightarrow \quad \textbf{FLAG}_{warmCldECM} = 1$$

with e.g. : $\Delta T = 5K$ for the uncertainty in the correct estimate of T.

4.3.2 Bayes results

The Bayes classification for tropospheric clouds works identically to the PSC cases. But due to the larger variability in the atmospheric background conditions (strong latitudinal variation in trace gas contribution) the training of the classifier needs to be separated for special conditions. Like illustrated in Table 4.4 a

restriction to latitude bands (e.g. polar, mid and equator latitudes various) can significantly improve the results. Other effective constraints are the optically thickness via the proxy CI-A and the reduction to scenarios with $TH \geq CBH$, which is equivalent to the restricted analysis of MIPAS spectra close to the CTH (first 1-2 cloudy sweeps in a profile).

cirrus/liquid	cir-mid	cir-psum	cir-tro	liq-mid	liq-psum	liq-tro
all	cir: 63.34 liq: 36.66	cir: 74.39 liq: 25.61	cir: 76.00 liq: 22.00	cir: 30.83 liq: 68.94	cir: 59.14 liq: 40.94	cir: 23.12 liq: 75.58
thin	cir: 67.92 liq: 32.08	cir: 77.95 liq: 22.05	cir: 20.33 liq: 36.50	cir: 13.96 liq: 85.85	cir: 57.49 liq: 42.06	cir: 3.84 liq: 61.86
thin/midl/ CBH	cir: **97.87** liq: 2.13	cir: 55.22 liq: 35.14	cir: 70.65 liq: 6.91	cir: 0.39 liq: **88.82**	cir: 20.93 liq: 65.61	cir: 9.37 liq: 49.09
thin/psum/ CBH	cir: 50.82 liq: 32.17	cir: **94.51** liq: 5.49	cir: 17.53 liq: 28.60	cir: 11.34 liq: 18.32	cir:6.67 liq: **93.33**	cir: 7.38 liq: 7.19

Table 4.4: Example of Bayes analyses for liquid and cirrus spectra of the CSDB. For training histograms (rows) and test spectra (columns) for all scenarios, optically thin events ($CI_A < 1.8$), thin ($1.8 < CI_A < 5$) scenarios for mid-latitude and tangent heights greater than the cloud base height (CBH), and the same extraction for polar summer scenarios. All analyses for classification histograms with 4 K bins.

For the application in the processor look-up tables of histograms for various latitude bands and CI-A ranges will be provided. Generally it is intended to restrict classification, macro and micro retrieval to first 2–3 cloud sweeps in a MIPAS scan.

4.4 Decision

The cloud type decision refers to the cloud decision (3.8) by using weighted sums of the different classification methods. For NAT and Ice clouds there a two methods

- Bayes: $\mathbf{FLAG}_{iceBayes}$, $\mathbf{FLAG}_{natBayes}$ and $\mathbf{FLAG}_{stsBayes}$
- TwoParameter-correlation classification: $\mathbf{FLAG}_{iceCorr}$ and $\mathbf{FLAG}_{natCorr}$

while for cirrus and liquid cloud type only the Bayes classification provides

- Bayes: $\mathbf{FLAG}_{cirBayes}$ and $\mathbf{FLAG}_{liqBayes}$

For each type the methods are added up with weights w:

$$\text{CONF}_{type} = (\mathbf{FLAG}_{typeBayes} \cdot w + \mathbf{FLAG}_{restrict} \cdot (-1)) + (\mathbf{FLAG}_{corr} \cdot w + \mathbf{FLAG}_{restrict} \cdot (-1))$$

The $\mathbf{FLAG}_{restrict}$ values will be provided by temperature criteria (see 4.3.1). The cloud will be identified as

Cloud Type	Decision Algorithm
unspecified	$\sum_i \text{CONF}_{type_i} \leq 0$)
ice	$\text{CONF}_{ice} > 0 \wedge \text{CONF}_{NAT} \leq 0 \wedge \text{CONF}_{STS} \leq 0$
nat	$\text{CONF}_{NAT} > 0 \wedge \text{CONF}_{ice} \leq 0 \wedge \text{CONF}_{STS} \leq 0$
sts	$\text{CONF}_{STS} > 0 \wedge \text{CONF}_{ice} \leq 0 \wedge \text{CONF}_{NAT} \leq 0$
cir	$\text{CONF}_{cir} > 0 \wedge \text{CONF}_{liq} \leq 0$
liq	$\text{CONF}_{liq} > 0 \wedge \text{CONF}_{cir} \leq 0$
mix (ice/nat)	$\text{CONF}_{ice} > 0 \wedge \text{CONF}_{NAT} > 0 \wedge \text{CONF}_{STS} \leq 0$
mix (ice/sts)	$\text{CONF}_{ice} > 0 \wedge \text{CONF}_{STS} > 0 \wedge \text{CONF}_{NAT} \leq 0$
mix (nat/sts)	$\text{CONF}_{STS} > 0 \wedge \text{CONF}_{NAT} > 0 \wedge \text{CONF}_{ice} \leq 0$
mix (ice/nat/sts)	$\text{CONF}_{ice} > 0 \wedge \text{CONF}_{NAT} > 0 \wedge \text{CONF}_{STS} > 0$

Table 4.5: Cloud type decision from measured confidences

5 Macro Retrieval

The three macroscopic parameters retrieved are

- z_c Cloud Top Height
- T_c Cloud Top Temperature
- k_c Cloud Extinction

The retrieval itself is a 3-stage process performed independently for each of a set of microwindows, followed by a 4th stage where the results from the different microwindows are combined and error estimates are produced.

The main part of this chapter follows the same structure:

5.1 Continuum Fit
5.2 Cloud Effective Fraction Retrieval
5.3 Cloud Parameter Retrieval
5.4 Combination and Error Estimation

In addition, there are appendices describing the following:

5.5.1 Microwindows
5.5.2 The Planck Function and its derivatives
5.5.3 Retrievable range of cloud extinction
5.5.4 Radiative transfer for pencil beams
5.5.5 Field-of-View convolution
5.5.6 Jacobians of the Cloud Effective Fraction
5.5.7 Handling optically thick clouds

5.1 Continuum Fit

The purpose of the continuum fit is to establish the continuum radiance contribution to each microwindow/tangent altitude, as distinct from the contributions due to molecular emission.

Starting with the pre-computed transmittance spectrum appropriate to the particular microwindow and tangent altitude, the points with the highest transmittance are used to determine the continuum radiance directly, as shown in Fig. 5.1

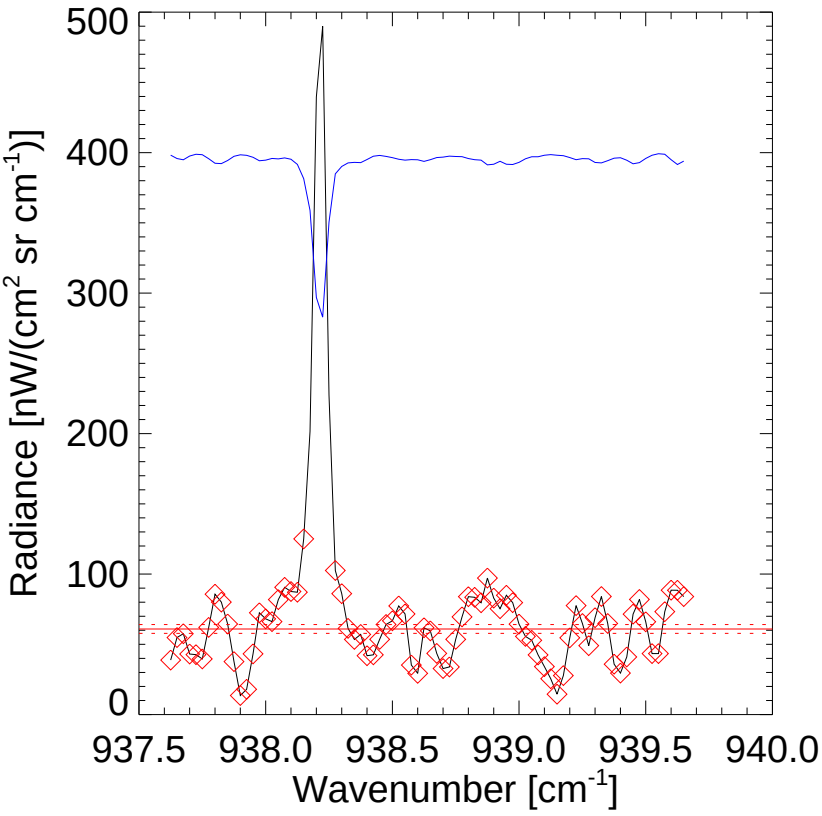

Figure 5.1: Illustration of the continuum fit. The black curve shows a measured MIPAS radiance at 18 km, the blue curve is the transmittance ($\times$400) for this altitude. The red diamonds indicate the points used to establish the continuum radiance (note the exclusion of points at the peak of the CO_2 line), shown by the horizontal red line ($\sim$61 nW), with an uncertainty ($\pm$4 nW) indicated by the dashed lines either side.

The criteria used to select points are

- Take all points $i = 1 \ldots n$ for which $\tau_i > \tau_{\min}$ (typically $\tau_{\min} = 0.95$)
- If $n < n_{\min}$ (typically $n_{\min}$=10% of total number of spectral points in microwindow), add further points i where $\tau_i < \tau_{\min}$ in decreasing order of transmittance until $n = n_{\min}$. This latter criterion is only likely to be necessary at low altitudes where the limb path transmittance of the atmospheric window regions is reduced.

A microwindow 3 cm^{-1} wide contains a total of 121 points at 0.025 cm^{-1} spectral sampling, 49 points at 0.0625 cm^{-1}.

The continuum radiance and uncertainty are then given simply by the mean and standard error of the n

spectral radiances $R_1 \ldots R_n$ used:

$$R_{\text{ctm}} = \frac{1}{n}\sum_i R_i \tag{5.1}$$

$$\sigma_{\text{ctm}} = \frac{1}{n}\sqrt{\sum_i (R_i - R_{\text{ctm}})^2} \tag{5.2}$$

Statistically, one would expect σ_{ctm} to have a value approaching NESR/$\sqrt{n}$, where NESR is the random noise associated with each spectral point. However, using the above calculation makes some allowance for additional uncertainties caused by residual (i.e., molecular) spectral structure in the points used. As a precaution against any statistical anomalies, σ_{ctm} is also limited to be $\geq$ NESR/$\sqrt{n}$.

5.2 CEF Retrieval

The purpose of the Cloud Effective Fraction (CEF) retrieval is twofold:

1. To identify the sweep containing the cloud-top
2. To determine the CEF for this sweep, which is used as an additional pseudo-measurement for the macroscopic parameter retrieval.

5.2.1 Cloud Effective Fraction

Cloud Effective Fraction, α is defined as

$$\alpha = \frac{R_{\text{ctm}}}{B_c} \tag{5.3}$$

where R_{ctm} is the continuum radiance in a spectral window (i.e., derived from the continuum fit described in the previous section), and $B_c(\nu, T_c)$ is the Cloud Top Radiance, which is the Planck function (see Appendix 5.5.2) for the microwindow central wavenumber ν and Cloud Top Temperature T_c.

CEF has a value between 0 and 1 and depends on two of the three macroscopic retrieval parameters, i.e., z_c, k_c. For thick cloud, it is simply the fraction of the field-of-view (FOV) which is filled by cloud, depending only on cloud top height. For thin cloud it is a more complicated function of cloud-extinction and cloud top height.

Since the CEF is determined from the spectral structure within a microwindow, it is regarded as independent of the vertical profile of continuum radiance. Both are used as inputs for the macroscopic parameter retrieval (5.3).

5.2.2 Atmospheric Model

This assumes a simple 'box' representation of the atmosphere where the height of the box represents the MIPAS field-of-view and the length is the atmospheric path. The upper part of the box is modelled as a cloud-free atmosphere with optically thin molecular absorption and the lower part is modelled as completely cloud-filled with continuum emission. Both the molecular and continuum emissions are assumed to be characterised

by the same Planck function B. A parameter α ($0 \leq \alpha \leq 1$) is the fraction of the box containing the cloud (i.e., the CEF). Thus the radiance is given by:

$$R(\nu) = (1-\alpha)B(1-\tau(\nu)) + \alpha B \tag{5.4}$$

where $\tau(\nu)$ is the pre-computed transmittance for this microwindow/altitude and B is the Planck function at cloud-top temperature computed at the spectral mid-point of the microwindow. The first term represents the molecular emission from the upper (cloud-free) part of the field-of-view and the second term the emission from the cloud-filled lower part.

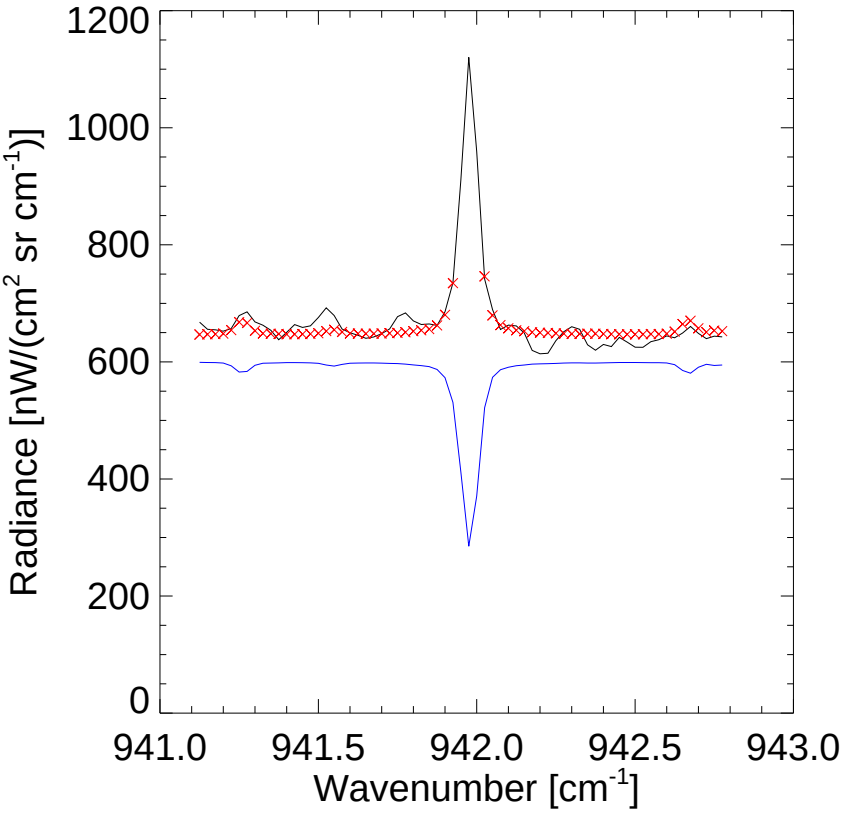

Figure 5.2: Illustration of the CEF retrieval. The black curve shows a measured MIPAS radiance at 15 km, the blue curve is the transmittance for this altitude ($\times$600). The red crosses indicate the fitted model radiance at the selected points. The CEF retrieved for this spectrum is 0.45 (and the retrieved cloud top radiance 1416 nW).

5.2.3 Selection of Spectral Points

This model relies on the spectral variation of transmittance to distinguish between molecular and continuum spectral contributions. However, if B is to represent the atmospheric temperature near the cloud top, it is still necessary to exclude points with strong absorption (low transmittance) so that only weak molecular lines are used. It is also assumed that the microwindow has been selected to contain only lines from well-mixed gases

(e.g. CO_2, see Appendix 5.5.1) so that the transmittance spectrum is reasonably representative of the actual atmospheric transmittance for this tangent height.

The criterion used to select points is

- Use all spectral points *i* such that $\tau(\nu_i) > \beta\tau_{\max}$ where $\tau_{\max}$ is the maximum transmittance of any spectral point in this microwindow/altitude and β is an adjustable parameter with typical value 0.75.

The measurement vector **y** is the set of m radiance values at the selected spectral points. The covariance is assumed to be diagonal $\mathbf{S}_y = \sigma_N^2\mathbf{I}$ where σ_N is the NESR. This is not strictly the case for apodised spectra but greatly simplifies the matrix inversion.

5.2.4 A Priori Estimate

The CEF retrieval is an optimal estimation scheme and therefore requires an *a priori* estimate of the state vector, which also serves as the initial guess.

The state vector **x** contains two elements (α, B). The *a priori* estimate **a** is defined as

$$\mathbf{a} = \begin{pmatrix} \alpha_{\text{apr}} \\ B_t \end{pmatrix} \qquad \mathbf{S}_a = \begin{pmatrix} \sigma_\alpha^2 & 0 \\ 0 & \sigma_{Bt}^2 \end{pmatrix} \tag{5.5}$$

where $\alpha_{\text{apr}} = 0$ is the estimate of the CEF with assumed uncertainty $\sigma_\alpha^2 = (1.0)^2$ (typically). The climatological (or other) temperature profile is used to determine $T(z_t)$ for the tangent point z_t, with associated uncertainty σ_{Tt}^2. This is then used to determine the corresponding Planck function B_t and associated variance $\sigma_{Bt}^2 = (dB/dT)^2\sigma_{Tt}^2$ (see Appendix 5.5.2).

5.2.5 Forward Model

The forward model spectrum **f** is computed from the state vector (α, B) (from Eq. 5.4):

$$f_i = (1-\alpha)B(1-\tau_i) + \alpha B \tag{5.6}$$

where subscript $i = 1 \ldots m$ denotes spectral points used. Elements of the $(m \times 2)$ Jacobian matrix **K** are also obtained from the differentials of this equation w.r.t. the state vector elements:

$$K_{i\alpha} = \frac{\partial f_i}{\partial \alpha} = B\tau_i \tag{5.7}$$

$$K_{iB} = \frac{\partial f_i}{\partial B} = 1 - \tau + \alpha\tau_i \tag{5.8}$$

5.2.6 Inverse Model

Starting with an initial guess $\mathbf{x}_0 = \mathbf{a}$, the retrieval is iterated in the standard form:

$$\mathbf{S}_{xj+1} = \left(\mathbf{K}_j^T \mathbf{S}_y^{-1} \mathbf{K}_j + \mathbf{S}_a^{-1}\right)^{-1} \tag{5.9}$$

$$\mathbf{x}_{j+1} = \mathbf{x}_j + \mathbf{S}_{xj+1}(\mathbf{K}_i^T \mathbf{S}_y^{-1}(\mathbf{y} - \mathbf{f}_j) - \mathbf{S}_a^{-1}(\mathbf{x}_j - \mathbf{a})) \tag{5.10}$$

where subscript j denotes iteration number.

After each iteration there is a check on the value of α (i.e., x_0) to limit it to the range $0 \leq \alpha \leq 1$.

For each iteration, the chi-squared statistic is computed from the measurement vector:

$$\chi_j^2 = (\mathbf{y} - \mathbf{f}_j)^T \mathbf{S}_y^{-1} (\mathbf{y} - \mathbf{f}_j) \tag{5.11}$$

and compared with the value from the previous iteration

$$\Delta\chi_j^2 = \chi_j^2 - \chi_{j-1}^2 \tag{5.12}$$

Convergence is reached when the absolute difference between successive values of χ^2 are less than some preset limit $\delta\chi^2$ (typically 0.01).

$$\left|\Delta\chi_j^2\right| < \delta\chi^2 \tag{5.13}$$

5.2.7 Outputs

The outputs of this section are the retrieved CEF, $\alpha = x_0$, and its associated standard deviation $\sigma_\alpha = \sqrt{(S_x)_{00}}$

If the retrieval does not converge within the maximum permitted number of iterations, the *a priori* value and covariance are substituted instead.

In principle, the updated cloud top radiance $B_c = x_1$ may also be useful for the macroscopic parameter retrieval, but because of the simplifications in this box model of the atmosphere, and because the *a priori* temperature itself may already be of relatively high quality (e.g., if derived from ECMWF), this is not used. This also avoids the complications of correlations between the CEF and the updated B_c that would have been introduced by this retrieval.

5.2.8 Cloud Top Identification

Starting with the highest altitude sweep, the cloud top is identified as occurring in the first sweep where the retrieved Cloud Effective Fraction α is greater than some threshold value: $\alpha > \alpha_C$ (typically $\alpha_C = 0.1$).

Note that the CEF retrieval model does not work reliably for completely cloud-filled fields-of-view when, for example, molecular lines may appear as absorption rather than emission features. Therefore having identified the highest spectrum which is partially cloud-filled, retrieved CEF values from lower altitude spectra are probably not useful.

5.3 Cloud Parameter Retrieval

The cloud parameter retrieval uses the continuum radiances (determined by the Continuum Fit, 5.1) and the CEF (determined by the CEF Retrieval, 5.2) to derive the three macroscopic parameters:

- z_c the Cloud Top Height [km]
- T_c the Cloud Top Temperature [K]
- k_c the Cloud Extinction [/km]

5.3.1 State Vector

Instead of directly retrieving Cloud Top Temperature, T_c, within each microwindow it is more convenient to deal with Cloud Top Radiance B_c, which is the Planck Function corresponding to T_c evaluated at the spectral mid-point of the microwindow (Appendix 5.5.2).

Also, rather than retrieve k_c (which can vary by orders of magnitude) it is more convenient to deal with the logarithm of this quantity, μ_c

$$\mu_c = \log_{10} k_c \tag{5.14}$$
$$k_c = 10^{\mu_c} \tag{5.15}$$

Derivatives with respect to μ_c and k_c are related as follows:

$$\frac{\partial}{\partial \mu_c} = \left(\frac{dk_c}{d\mu_c}\right)\frac{\partial}{\partial k_c} = k_c \ln(10)\frac{\partial}{\partial k_c} \tag{5.16}$$

Thus the state vector **x** contains

$$\mathbf{x} \equiv \begin{pmatrix} z_c \\ B_c \\ \mu_c \end{pmatrix} \tag{5.17}$$

5.3.2 Measurement Vector

The cloud parameter retrieval uses a measurement vector **y** containing

- R_c the continuum radiance from the sweep identified as containing the cloud-top
- R_b the continuum radiance from the sweep immediately below
- R_t the continuum radiance from the sweep immediately above
- α the Cloud Effective Fraction

The first three are generated from the continuum fit (5.1) and the last from the CEF retrieval (5.2), along with their associated 1σ uncertainties which are squared to form the diagonal elements of the measurement covariance matrix (assuming no correlations).

R_c and α will always be present, although, in the case of α, possibly with just the *a priori* value of zero. If the cloud top is identified as occurring in the lowest sweep of a scan then R_b will be absent. Also possible, but less likely, is that the cloud top is identified in the highest sweep, in which case R_t will be absent. Either or both of R_b and R_c could also be missing due to their spectra being flagged as corrupt.

Thus the measurement vector contains **y** contains a minimum of two elements (R_c, α) and a maximum of four elements (R_t, R_c, R_b, α).

5.3.3 A Priori Estimate

Since this is an optimal estimation scheme, *a priori* estimates, and associated uncertainties, are required for each element of the state vector (Eq. 5.17). These are also used as the initial values for the iterative retrieval.

Having identified the sweep containing the cloud top, the engineering tangent height of this sweep z_t also acts as an *a priori* estimate of z_c with an assumed uncertainty of σ^2_{za} representing

$$\sigma^2_{za} = \langle (z_c - z_t)^2 \rangle \tag{5.18}$$

Since the cloud top has been identified as falling within this field-of-view (height $\sim$3 km) a reasonable value for σ^2_{za} is (1 km)^2.

There is already (from 5.2.4) an *a priori* estimate of the radiance B_t (and uncertainty σ_{Bt}) associated with the tangent altitude z_t. Since z_t is the *a priori* estimate of the cloud top height, then B_t is the *a priori* estimate of the cloud top radiance. However, when constructing the covariance matrix for these quantities it is also necessary to consider the difference between these quantities and the actual cloud top parameters B_c, z_c:

$$B_c = B_t + \frac{dB}{dz}(z_c - z_t) \tag{5.19}$$

where the vertical gradient of the Planck function dB/dz is obtained from the climatological temperature profile dT/dz

$$\left(\frac{dB}{dz}\right) = \left(\frac{dB}{dT}\right)\left(\frac{dT}{dz}\right) \tag{5.20}$$

(see Appendix 5.5.2 for dB/dT).

Thus the variance of the *a priori* cloud top radiance estimate is given by

$$\sigma^2_{Ba} = \langle (B_c - B_t)^2 \rangle = \sigma^2_{Bt} + \left(\frac{dB}{dz}\right)^2 \sigma^2_{za} \tag{5.21}$$

and there will also be a correlation term

$$\sigma_{Ba}\sigma_{za} = \langle (B_c - B_t)(z_c - z_t) \rangle = \left(\frac{dB}{dz}\right) \sigma^2_{za} \tag{5.22}$$

Thus, the *a priori* vector and covariance are given by

$$\mathbf{a} = \begin{pmatrix} z_t \\ B_t \\ \mu_a \end{pmatrix} \tag{5.23}$$

$$\mathbf{S}_a = \begin{pmatrix} \sigma_{za}^2 & \sigma_{Ba}\sigma_{za} & 0 \\ \sigma_{Ba}\sigma_{za} & \sigma_{Ba}^2 & 0 \\ 0 & 0 & \sigma_{\mu a}^2 \end{pmatrix} \tag{5.24}$$

Finally, for the extinction element μ, it is not possible to construct any meaningful *a priori* estimate. However we can estimate the range of retrievable values as $-4 \leq \mu_c \leq -1$ (Appendix 5.5.3) so setting a *a priori* value of $\mu_a = -2.5$ with an uncertainty $\sigma_{\mu a}^2 = (0.5)^2$ seems reasonable. Alternatively we could assume a thick cloud and set $\mu_a = -1.0$. In this case (Appendix 5.5.7), the forward model becomes insensitive to cloud extinction and the retrieval is effectively reduced to two parameters: B_c and z_c. Since the *a priori* estimate is also used as the initial guess, significantly different answers may be obtained depending on which value is used.

5.3.4 Forward Model

The forward model **f** simulates the measurements (radiances and CEF) expected from a cloud described by the three macroscopic parameters given by the state vector for the given viewing geometry (tangent heights and FOV shape)

$$\mathbf{f} \equiv \mathbf{f}(\mathbf{x}) \tag{5.25}$$

It also calculates the derivatives of these quantities for the Jacobian matrix

$$\mathbf{K} \equiv \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \tag{5.26}$$

The mathematical details of the forward model are described in Appendices 5.5.4–5.5.6.

5.3.5 Inverse Model

The retrieval is an iterative optimal estimation scheme, similar to that used for the CEF retrieval, but with the addition of a Levenberg-Marquardt parameter γ (5.3.6) to aid convergence.

Starting with an initial guess $\mathbf{x}_0 = \mathbf{a}$, the retrieval is iterated in the standard form:

$$\mathbf{S}_{x\,j+1} = \left(\mathbf{K}_j^T \mathbf{S}_y^{-1} \mathbf{K}_j + (1+\gamma)\mathbf{S}_a^{-1}\right)^{-1} \tag{5.27}$$

$$\mathbf{x}_{j+1} = \mathbf{x}_j + \mathbf{S}_{x\,j+1}(\mathbf{K}_i^T \mathbf{S}_y^{-1}(\mathbf{y} - \mathbf{f}_j) - \mathbf{S}_a^{-1}(\mathbf{x}_j - \mathbf{a})) \tag{5.28}$$

where subscript j denotes iteration number.

5.3.6 Convergence

The Levenberg-Marquardt parameter γ is initially set to a small value $\gamma_{\min}$ (=0.01) so that on the first iteration, Eqs. (5.27),(5.28) approximate the standard optimal estimation update (e.g., Eqs. 5.9,5.10).

After each iteration j, a χ^2 statistic is calculated for the measurement vector:

$$\chi_j^2 = (\mathbf{y}-\mathbf{f}_j)^T \mathbf{S}_y^{-1}(\mathbf{y}-\mathbf{f}_j) \tag{5.29}$$

This is compared to the previous iteration

$$\Delta\chi_j^2 = \chi_j^2 - \chi_{j-1}^2 \tag{5.30}$$

The initial value χ_0^2 is calculated from the forward model $\mathbf{f}$ applied to the initial guess value $\mathbf{x}_0 = \mathbf{a}$, i.e., $\mathbf{f}_0 = \mathbf{f}(\mathbf{a})$.

The change in χ^2 is then compared to some preset limit, $\delta\chi^2$ (typically 0.01):

- $|\Delta\chi_j^2| < \delta\chi^2$ and $\gamma = \gamma_{\min}$: stop iterations, retrieval converged; otherwise apply remaining tests
- $\Delta\chi_j^2 < 0$: decrease $\gamma \rightarrow \max(\gamma/10, \gamma_{\min})$ and next iteration
- $\Delta\chi_j^2 > 0$ and $\gamma < \gamma_{\max}$: increase $\gamma \rightarrow \gamma \times 10$ and repeat last iteration
- $\Delta\chi_j^2 > 0$ and $\gamma = \gamma_{\max}$: stop iterations, retrieval diverged.

There is also a limit on the maximum number of iterations. If this is reached then the retrieval is stopped and flagged as not converged (same as if it diverges in the above list).

5.3.7 Cloud Top Temperature

The retrieval state vector contains Cloud Top Radiance B_c but this depends partly on the spectral position of the microwindow rather than just the cloud itself. So before results from different microwindows can be combined it is necessary to convert this to Cloud Top Temperature (see Appendix 5.5.2):

$$T_c = \frac{c_2 \nu}{\ln(1 + c_1 \nu^3 / B_c)} \tag{5.31}$$

The covariance matrix also has to be converted. In linearised form the transformation of the state vector can be represented by

$$\mathbf{x}' = \mathbf{G}\mathbf{x} \tag{5.32}$$

where **G** is the diagonal matrix with elements $(1, dB/dT, 1)$. Hence the appropriate conversion for the covariance is given by

$$\mathbf{S}_x' = \mathbf{G}\mathbf{S}_x\mathbf{G}^T \tag{5.33}$$

Effectively this means scaling elements of $\mathbf{S}_x$ as follows (assuming x_2 is the state vector element for cloud top radiance/temperature))

$$S_{ij} \times (dB/dT)^2 \quad \text{if } i = j = 2 \tag{5.34}$$

$$S_{ij} \times 1 \quad \text{if } i \neq 2 \text{ and } j \neq 2 \tag{5.35}$$

$$S_{ij} \times (dB/dT) \quad \text{elsewhere} \tag{5.36}$$

5.3.8 Outputs

The retrieval is run three times with different measurements/initial guess values:

1. If available (i.e., cloud top not located in the lowest sweep in the scan) using the measurement from the sweep below the cloud top, R_b with *a priori* $\log_{10}$(extinction), $\mu_a = -2.5$. This is the full 3 parameter retrieval (z_c, T_c, μ_c) from 3 measurements (R_c, R_b, α) (plus the nominally zero radiance measurement R_t from the sweep above the cloud top).
2. As (1) but setting $\mu_{ca} = -1.0$, i.e., 'thick cloud' assumption (Appendix 5.5.7) with effectively 2 parameters (z_c, T_c) retrieved from 3 measurements (R_c, R_b, α).
3. As (2) but without R_b, i.e., 'thick cloud' assumption and retrieving two parameters (z_c, T_c) from only one sweep using two measurements (R_c, α). This relies on the (idealised) CEF retrieval in order to separate the two parameters.

In each case, the outputs are the 3-element retrieved state vector and the 3×3 covariance matrix, plus a flag indicating whether or not the retrieval converged. Note that even when it is effectively a 2 parameter retrieval the output contains all 3 elements to simplify coding, even though the 'retrieved' extinction and covariance will essentially just reproduce the *a priori* values.

5.3.9 Cloud Top Pressure

In addition to the three retrieved parameters a cloud top pressure value is interpolated from the corresponding nearest L1B data neighbours of the retrieved cloud top height value and the ecmwf pressure profile. Interpolation is performed in log space.

5.4 Combination and Error Estimation

For each microwindow, three alternative retrievals (and associated covariances) are available (5.3.8). This section deals with the combination of results to determine the best estimate of the output parameters and the estimation of uncertainties in these quantities.

5.4.1 Combining microwindows

If there are N independent estimates of a vector $\mathbf{x}_i$ $(i = 1 \ldots N)$ each with associated covariance $\mathbf{S}_{xi}$, then the covariance and mean are given by the vector equivalent of weighting by inverse variances:

$$\hat{\mathbf{S}}_x^{-1} = \sum_{i=1}^{N} {\mathbf{S}_{xi}}^{-1} \tag{5.37}$$

$$\hat{\mathbf{x}} = \hat{\mathbf{S}}_x \left(\sum_{i=1}^{N} {\mathbf{S}_{xi}}^{-1} \mathbf{x}_i \right) \tag{5.38}$$

Separate combined estimates are obtained in this way from each of the three retrievals using, in each case, just the microwindows which have converged for that retrieval.

5.4.2 Spike Test

The combination process is also a convenient point at which to perform a spike test to remove any microwindow results which fall well outside the expected distribution.

For each contributing microwindow a χ^2 statistic is evaluated:

$$\chi_i^2 = (\mathbf{x}_i - \hat{\mathbf{x}})^T \hat{\mathbf{S}}_x^{-1} (\mathbf{x}_i - \hat{\mathbf{x}}) \tag{5.39}$$

If the microwindow which has the maximum χ_i^2 exceeds the average χ_i^2 by some specified factor (typically 2), then that microwindow is removed and the remaining $N-1$ microwindow results recombined.

This process is repeated until all remaining microwindows pass the spike test (which may mean just one microwindow).

5.4.3 Selecting Retrieval

Following the spike test, up to three different combined retrieval results are available, corresponding to the three different *a priori*/measurement set-ups described in 5.3.8.

The results are tested in the given order and the selected result is the first one for which the number of microwindows which remain after the spike test (i.e., which have also converged) exceeds some defined minimum value, typically 5 (i.e. half the total number of microwindows used).

The reasoning is that the 'full' 3 parameter, 3 measurement retrieval (1) is the most useful but also the least likely to converge, or to converge to a consistent solution for all microwindows. The other retrievals are increasing more robust in that they are either more constrained (2) or remove the assumption that adjacent sweeps view the same cloud (3), but retrieve just two parameters (and possibly just one in the case of (3)).

5.4.4 Error Characterisation

In principle, the combined covariance $\hat{\mathbf{S}}_x$ (Eq. 5.37) describes the uncertainty in each retrieved parameter and their correlations. However this will be an underestimate since it takes no account of any shortcomings in the simplistic forward model.

An alternative would be to use the actual variation in the results from each microwindow which contribute to the mean. However, with relatively few statistical samples it would be difficult to obtain reliable results, especially if correlations are also required.

The adopted solution is to use the correlations from the combined covariance but to scale the errors upwards, if necessary, to match the observed scatter. This is achieved by constructing a 3-element vector of scale

factors, **e**:

$$e_i = \max(\hat{\sigma}_i, \mathrm{SD}_i) \tag{5.40}$$

where σ_i is the square root of diagonal element ii in matrix $\hat{\mathbf{S}}_x$ (i.e., the uncertainty in parameter x_i according to the covariance matrix) and SD_i is the actual standard deviation of the parameter x_i from the different microwindow results.

The retrieval covariance is then 'adjusted' to produce the final covariance as follows

$$\hat{\mathbf{S}}'_x = \mathbf{e}^T \hat{\mathbf{S}}_x \mathbf{e} \tag{5.41}$$

Thus the resulting error covariance matrix will have diagonal elements which represent the variation in results from the different microwindows (or the originally predicted uncertainty, if larger) but maintain the correlations from the original covariance matrix.

5.4.5 Quality Flags

In addition to the three macroscopic parameters and their covariance, the following overall quality flag is also output.

-1 No cloud detected within scan range. No retrieval output

0 Cloud detected but insufficient converged retrievals to give meaningful result. No retrieval output

1 Results obtained from single sweep above cloud top, thick cloud assumption (z_c, T_c, but possibly highly correlated)

2 Results obtained from two sweeps (above and below cloud top), thick cloud assumption (z_c, T_c)

3 Results obtained from two sweeps, cloud extinction also retrieved (z_c, T_c, k_c)

In addition to that, there are three equivalents to the presented flags $1, 2, 3$ which are $-2, -3, -4$. They indicate that the retrieved CTH is outside an interval of the smallest/highest CEF detected CTH $\pm 2km$ (which represents the field of view) over all converged micro windows and thus should be handled with caution.

-2 Results obtained from single sweep above cloud top, thick cloud assumption (z_c, T_c, but possibly highly correlated), but retrieved CTH outside FOV

-3 Results obtained from two sweeps (above and below cloud top), thick cloud assumption (z_c, T_c), but retrieved CTH outside FOV

-4 Results obtained from two sweeps (above and below cloud top), cloud extinction also retrieved (z_c, T_c, k_c), but retrieved CTH outside FOV

Although no macroscopic parameters are output for flags -1 and 0, the CEF for each sweep may still be useful. If the negative flags $-2, -3$ or -4 are set for one scan then the result shall not be used for the summary result of cloud detection.

5.5 Appendix

5.5.1 Microwindows

The retrieval uses a pre-defined set of microwindows as well as associated transmittance spectra.

Ideally these microwindows are located in an optically thin region of the spectrum (so that the continuum contribution of any cloud can be determined) and where the main molecular absorbers are well-mixed gases (to help with the joint temperature/Cloud Effective Fraction determination). The microwindows initially selected for this study are listed in Table 5.1 and plotted in Fig. 5.3.

Table 5.1: List of microwindows for cloud macroscopic parameter retrieval from full resolution MIPAS spectra (see also Fig. 5.3)

MW#	Wavenumber Range [cm^{-1}]	
1	937.6250	940.6250
2	941.1250	944.1250
3	944.5500	947.5500
4	955.7750	958.7750
5	948.6250	951.1750
6	935.9500	937.6000
7	934.4500	935.8750
8	953.4750	955.0500
9	951.9000	953.2500
10	958.8000	960.8250

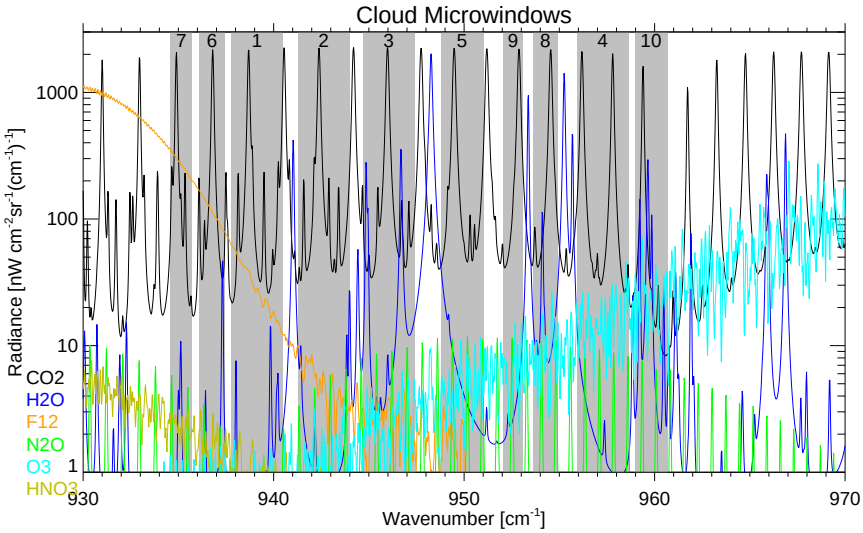

Figure 5.3: The shaded areas and numbers show the 10 microwindows (in priority order) selected for the macroscopic parameter retrieval using full-resolution MIPAS spectra (more details in Table 5.1). The modelled spectra for a tangent height of 9 km for the major emitting molecules in this spectral region are also shown.

From the figure, it is evident that the microwindows are all in the 10 μm atmospheric window, and each includes one or two CO_2 lines (actually from the laser band, but it is assumed that non-LTE effects can be ignored at these low altitudes) but excludes the (highly variable) H_2O lines. These microwindows were actually selected by an adapted, and simplified, version of the MWMAKE algorithm developed for MIPAS pT and VMR retrievals. However, an equally good subjective selection could probably have been made by eye — the microwindow selection algorithm is not regarded as essential to this method.

It is assumed that, unlike in the MIPAS operational retrievals, there are no spectral masks or altitude ranges associated with the microwindows. However, if required, spectral points can be excluded from use by setting the transmittance spectra to zero at those points.

There is no fundamental reason why the microwindows all have to be in the same spectral region, however there is an assumption that the same cloud extinction value is retrieved from all microwindows so having them closely grouped is preferable.

5.5.2 Planck Function

The Planck function B, expressed in terms of wavenumber ν and temperature T, is:

$$B(\nu,T) = \frac{c_1\nu^3}{\exp(c_2\nu/T)-1} \tag{5.42}$$

where $c_1 = 1.191\times10^{-3}$ nW/(cm^2 sr (cm^{-1})^4) and $c_2 = 1.439$ K/cm^{-1}.

This can be inverted to give:

$$T(\nu,B) = \frac{c_2\nu}{\ln(1+c_1\nu^3/B)} \tag{5.43}$$

Also useful are various forms of the derivative:

$$\frac{dB}{dT} = \frac{c_1c_2\nu^4\exp(c_2\nu/T)}{(\exp(c_2\nu/T)-1)^2T^2} \tag{5.44}$$

$$= \left(\frac{B}{T^2}\right)\frac{c_2\nu}{(1-\exp(-c_2\nu/T))} \tag{5.45}$$

$$= \frac{B^2b(\ln b)^2}{c_1c_2\nu^4} \tag{5.46}$$

where $b = 1+c_1\nu^3/B$.

5.5.3 Cloud Extinction range

To estimate the range of cloud extinction values that may be retrieved consider the limits of optically thin and optically thick clouds.

In the optically thin limit, a completely cloud-filled FOV has radiance

$$R_c = B_c(1-\exp(-k_cx)) \simeq B_ck_cx \tag{5.47}$$

We can also express the CEF in the optically thin limit as

$$\alpha = \frac{R_c}{B_c} \simeq k_c x \tag{5.48}$$

So, assuming a path length of $x = 300$ km, a minimum detection limit of $\alpha = 0.1$ corresponds to a minimum extinction value of

$$k_c = \frac{0.1}{300} \simeq 3 \times 10^{-4} \text{ km}^{-1} \tag{5.49}$$

i.e., the retrieved $\log_{10} k_c = \mu_c \geq -3.5$.

In the optically thick limit, if we assume that the extinction is indistinguishable from infinity once the path transmittance reaches 1%,

$$\exp(-k_c x) = 0.01 \tag{5.50}$$

$$k_c = -\frac{\ln(0.01)}{300} \tag{5.51}$$

$$= 1.5 \times 10^{-2} \text{ km}^{-1} \tag{5.52}$$

giving $\mu_c \leq -1.8$.

Thus we expect to retrieve reasonable values of $\log_{10} k_c$ in the approximate range $-4 \leq \mu_c \leq -1$.

5.5.4 Pencil Beams

5.5.4.1 Isothermal Cloud

First, consider the case of a homogeneous isothermal cloud with cloud top altitude z_c, cloud top radiance B_c and extinction k_c, Fig. 5.4

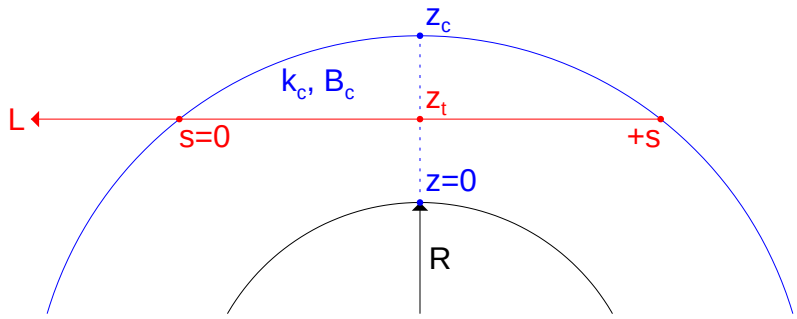

Figure 5.4: Geometry for a homogeneous cloud with cloud top height z_c, and a pencil beam viewing tangent height z_t. See text for other notation.

The radiance of the pencil beam is given by L

$$L = B_c(1 - \tau) \tag{5.53}$$

where τ is the transmittance of the path. Derivatives of this radiance L with respect to the three retrieved parameters are as follows:

$$\frac{\partial L}{\partial B_c} = 1-\tau \tag{5.54}$$

$$\frac{\partial L}{\partial z_c} = -B_c\frac{d\tau}{dz_c} \tag{5.55}$$

$$\frac{\partial L}{\partial k_c} = -B_c\frac{d\tau}{dk_c} \tag{5.56}$$

The path transmittance is given by

$$\tau = \exp(-k_c s) \tag{5.57}$$

where the distance s of a pencil beam within the cloud at tangent height z_t is given by Pythagoras' theorem, assuming that the earth radius $R \gg z_c, z_t$,

$$s = 2\sqrt{2R(z_c-z_t)} \tag{5.58}$$

$$\frac{ds}{dz_c} = \frac{(2R)^{\frac{1}{2}}}{(z_c-z_t)^{\frac{1}{2}}} \tag{5.59}$$

Hence derivatives of τ w.r.t. retrieved parameters are

$$\frac{d\tau}{dk_c} = -s\tau = -2(2R)^{\frac{1}{2}}(z_c-z_t)^{\frac{1}{2}}\tau \tag{5.60}$$

$$\frac{d\tau}{dz_c} = -k_c\frac{ds}{dz_c}\tau = -k\frac{(2R)^{\frac{1}{2}}}{(z_c-z_t)^{\frac{1}{2}}}\tau \tag{5.61}$$

5.5.4.2 Vertical Temperature Gradient

Next consider the case of a known vertical temperature gradient within the cloud, such that the Planck function at altitude z is related to the cloud top Planck function B_c by a constant b (negative in the troposphere)

$$B(z) = B_c + b(z-z_c) \tag{5.62}$$

The pencil beam radiance now has to allow for the variation of B with z along the path

$$L = -\int B\frac{d\tau}{ds}ds \tag{5.63}$$

where the path coordinate s is defined as zero at the cloud top nearest the satellite and $+s$ at the far end. (Note that Eq. (5.53) is the simplification of the above if $B = B_c$ is taken as constant). Making a coordinate transformation from (s,z) to (x,y) so that $x = 0, y = 0$ defines the tangent point (Fig. 5.5)

$$x^2 = 2Ry \tag{5.64}$$

The transmittance is defined in terms of the new path coordinate x as

$$\tau = \exp -(x+x_c) \tag{5.65}$$

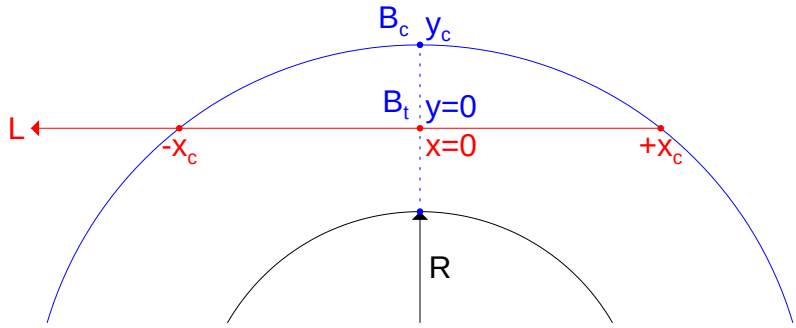

Figure 5.5: As Fig. 5.4 but with B varying with altitude and coordinates (x,y) replacing (s,z).

where $x = -x_c$ at the cloud top nearest the satellite. From Pythagoras' Theorem

$$x_c^2 \simeq 2R(z_c - z_t) = 2Ry_c \tag{5.66}$$

The Planck function variation with altitude becomes

$$B(y) = B_t + by \tag{5.67}$$

where B_t, the Planck function at the tangent point, is given by

$$B_t = B_c + b(z_t - z_c) \tag{5.68}$$

This can be rewritten in terms of the path coordinate as

$$B(x) = B_t + bx^2/(2R) \tag{5.69}$$

The radiative transfer equation (Eq. 5.63) then becomes

$$L = \int_{-x_c}^{+x_c} \left(B_t + \frac{bx^2}{2R}\right) k_c \exp -k_c(x + x_c)\, dx \tag{5.70}$$

which has the solution

$$L = \left(B_c + \frac{b}{Rk_c^2}\right)(1-\tau) - \frac{bs}{2Rk_c}(1+\tau) \tag{5.71}$$

where the total path length $s = 2x_c$ and $\tau = \exp(-k_c s)$.

Note that in the limit of zero vertical gradient, $b = 0$, Eq. (5.71) simplifies to that for the isothermal cloud, Eq. (5.53), as expected.

While Eq. (5.71) represents the exact solution, the appearance of k_c, k_c^2 in the denominator may lead to numerical instabilities in the thin cloud limit ($k_c \to 0$). Therefore a simpler, more robust form will be required.

5.5.4.3 Asymptotic Limits

In the case of thick cloud, $\tau = 0, k_c \sim \infty$ and Eq. (5.71) converges to the trivial solution

$$L(\tau \to 0) = B_c \tag{5.72}$$

i.e., the cloud acts as a black body at its cloud top radiance, and any vertical temperature structure (b) within the cloud is irrelevant.

The radiance in the thin cloud limit can be obtained by expanding $\tau = \exp(-k_c s)$ as a polynomial (third order required because of the $1/k_c^2$ in the first bracket of Eq. (5.71)). This gives

$$L(\tau \to 1) = \left(B_c - \frac{b}{6} \left(\frac{s^2}{2R} \right) \right) k_c s \tag{5.73}$$

Making substitutions for x_c (=$s/2$) from Eq. (5.66) and for b from Eq. (5.62), this can be rewritten as:

$$L(\tau \to 1) = \left(\frac{1}{3} B_c + \frac{2}{3} B_t \right) k_c s \tag{5.74}$$

which makes it more obvious that the effective emitting temperature of a thin cloud approaches that of the point $1/3$ distance (in altitude) from the tangent point to the cloud top.[1]

5.5.4.4 Computational Form

To avoid numerical instabilities in the thin cloud limit, it will be assumed that Eq. (5.71), can be well approximated as as a linear function of τ between the thick (Eq. 5.72) and thin (Eq. 5.74) cloud asymptotic limits:

$$L = \left(B_c + \frac{2}{3} (B_t - B_c) \tau \right) (1 - \tau) \tag{5.75}$$

Expressing B_t (Eq. 5.68) in terms of retrieved parameters, this becomes

$$L = \left(B_c + \frac{2}{3} b (z_t - z_c) \tau \right) (1 - \tau) \tag{5.76}$$

The derivatives with respect to the retrieved macroscopic parameters are then as follows:

$$\frac{\partial L}{\partial B_c} = 1 - \tau \tag{5.77}$$

$$\frac{\partial L}{\partial z_c} = -\left[B_c - \frac{2}{3} b (z_t - z_c)(1 - 2\tau) \right] \frac{d\tau}{dz_c} - \frac{2}{3} b \tau (1 - \tau) \tag{5.78}$$

$$\frac{\partial L}{\partial k_c} = -\left[B_c - \frac{2}{3} b (z_t - z_c)(1 - 2\tau) \right] \frac{d\tau}{dk_c} \tag{5.79}$$

which, as expected, converge to the equivalent expressions for isothermal clouds (Eqs. 5.54–5.55) in the limit of zero vertical temperature gradient ($b = 0$).

[1] This particular result, for the effective emitting radiance in the thin cloud limit, could also be obtained more simply by taken the path-weighted average of B: $\int B dx / s$

5.5.5 FOV Convolution

5.5.5.1 Numerical Form

The measured radiance $R(z_t)$ at a nominal tangent height z_t (i.e., the tangent height of the centre of the FOV) represents a convolution of the pencil beam radiances $L(z)$ with the instrument field-of-view function $\Phi(z-z_t)$

$$R(z_t) = \frac{\int L(z)\Phi(z-z_t)\,dz}{\int \Phi(z-z_t)\,dz} \tag{5.80}$$

In practice, the instrument FOV function Φ is tabulated at a series of points $\Phi_i \equiv \Phi(z_i - z_t)$, which also define the tangent paths z_i for which the pencil-bean radiances are calculated $L_i \equiv L(z_i)$, as shown in Fig. 5.6. The integral is then approximated by a summation

$$R(z_t) \simeq \frac{\sum \phi_i L_i}{\sum \phi_i} \tag{5.81}$$

where the relationship between the pencil beam weights ϕ_i and the tabulated FOV function Φ_i depends on the assumed behaviour of the radiance field between tabulation points (actually also on the behaviour of the FOV function Φ, but this is always assumed to vary linearly).

5.5.5.2 Smoothly Varying Radiance

In an interval where both the upper z_{i+1} and lower z_i paths intersect the cloud, it is assumed that the radiance $L(z)$ varies linearly with altitude:

$$L(z) = L_i + \left(\frac{\Delta L}{\Delta z}\right)(z - z_i) \tag{5.82}$$

where $\Delta L \equiv (L_{i+1} - L_i)$ and similarly for Δz.

The FOV function Φ is also assumed to vary linearly:

$$\Phi(z) = \Phi_i + \left(\frac{\Delta \Phi}{\Delta z}\right)(z - z_i) \tag{5.83}$$

Integrating the product of these two terms over the interval $i : i+1$:

$$\delta R_i = \int_i^{i+1} L\Phi\,dz \tag{5.84}$$

$$= \frac{1}{6}\Big(L_i\,(2\Phi_i + \Phi_{i+1}) + L_{i+1}\,(\Phi_i + 2\Phi_{i+1})\Big)\Delta z \tag{5.85}$$

Where this linear radiance approximation is used either side of a point i, the weight for that pencil beam radiance becomes

$$\phi_i = \frac{1}{6}\Big((\Phi_{i-1} + 2\Phi_i)(z_i - z_{i-1}) + (\Phi_{i+1} + 2\Phi_i)(z_{i+1} - z_i)\Big) \tag{5.86}$$

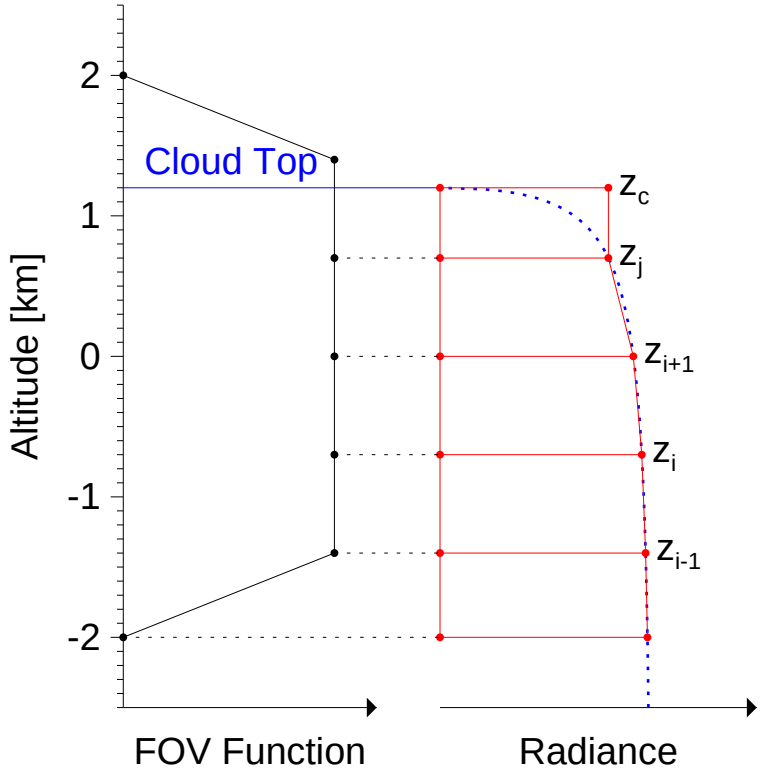

Figure 5.6: Field-of-View convolution of pencil beams. The FOV function is plotted on the left hand side, with the dots representing the tabulated point. The cloud radiance profile is shown by the blue dotted line on the right (calculated for an extinction of 0.01/km). The red trapezia represent the areas used for integration.

At the limits of the tabulation, $\Phi_1, \Phi_n = 0$ (therefore also $\Phi_0, \Phi_{n+1} = 0$), the weights for pencil beam radiances at the end points simplify to

$$\phi_1 = \Phi_2(z_2 - z_1) \quad (5.87)$$

$$\phi_n = \Phi_{n-1}(z_n - z_{n-1}) \quad (5.88)$$

5.5.5.3 Cloud Top

While Eq. (5.85) is appropriate for the integration over regions of the field-of-view where the continuum radiance varies smoothly with altitude, it is not a good approximation for the interval which spans the cloud-top.

A better approximation for this particular interval can be obtained if it is assumed that the radiance behaves as a step function[2]: zero above the cloud top, and a fixed radiance below with a discontinuity at the cloud top height, as shown in Fig. 5.6 as the top red rectangle.

[2] Strictly, the step function represents the limit for thick cloud. For thinner cloud, as plotted in the figure, it can be seen that the blue dotted line departs significantly in this region, tending to a linear decrease to zero at the cloud top altitude. However the choice is to accurately model the cloud top in the optically thick regime.

If the cloud-top z_c lies between two tabulation points $z_j \leq z_c < z_{j+1}$ then, assuming the FOV function varies linearly with z as before (Eq. 5.83), the integral for this segment becomes

$$\begin{aligned} \delta R_j &= \int_{z_j}^{z_c} L_j \Phi \, dz && (5.89) \\ &= L_j \Phi_j (z_c - z_j) + \\ &\quad L_j \left(\Phi_{j+1} - \Phi_j\right) \frac{(z_c - z_j)^2}{2\Delta z} && (5.90) \end{aligned}$$

Combining this with Eq. (5.85) representing the contribution of the pencil beam to the interval $(j-1:j)$, the overall weight for pencil beam L_j is given by

$$\begin{aligned} \phi_j &= \Phi_j (z_c - z_j) \\ &\quad + \left(\Phi_{j+1} - \Phi_j\right) \frac{(z_c - z_j)^2}{2\Delta z} \\ &\quad + \frac{1}{6} \left(\Phi_{j-1} + 2\Phi_j\right) (z_j - z_{j-1}) && (5.91) \end{aligned}$$

Since the radiance above the cloud top is zero, there is no need to calculate a weight ϕ_{j+1} for the upper altitude for this interval in the FOV function.

5.5.5.4 FOV Derivatives

The FOV-convolved radiance is given by Eq. (5.81). For the retrieval it is necessary also to have the Jacobians, i.e., derivatives of this expression with respect to retrieved parameters. For most terms in the summation this is simply the derivative of the pencil beam radiances L_i, with weights ϕ_i treated as constant since they are independent of the retrieved parameters. However, this is not true of the weight ϕ_j (Eq. 5.91) of the pencil beam just below the cloud top, which depends on the retrieved cloud-top height, z_c. Differentiating Eq. (5.91) gives

$$\frac{\partial \phi_j}{\partial z_c} = \Phi_j + (\Phi_{j+1} - \Phi_j) \frac{(z_c - z_j)}{\Delta z} \tag{5.92}$$

Thus the derivative of the FOV-convolved radiance w.r.t. cloud-top height becomes

$$\begin{aligned} \frac{\partial R(z_t)}{\partial z_c} &= \frac{1}{\sum_i \phi_i} \sum_i \frac{\partial (\phi_i L_i)}{\partial z_c} && (5.93) \\ &= \frac{1}{\sum_{i=1}^{n} \phi_i} \left(\sum_{i=1}^{j} \phi_i \frac{\partial L_i}{\partial z_c} + L_j \frac{\partial \phi_j}{\partial z_c} \right) && (5.94) \end{aligned}$$

Note that since the pencil beam radiances (and derivatives) above the cloud top ($i > j$) are zero, the summation in the numerator need only be evaluated for $i = 1 \ldots j$, however the full summation $i = 1 \ldots n$ is still required for the normalisation in the denominator.

5.5.6 CEF Derivatives

The Cloud Effective Fraction α is used as a 'measurement' in the macroscopic parameter retrieval, therefore derivatives are required w.r.t. the retrieved parameters.

The CEF is defined (Eq. 5.3) as the ratio of the continuum radiance R_c (in the spectrum containing the cloud-top) to the cloud top radiance B_c:

$$\alpha = \frac{R_c}{B_c} \tag{5.95}$$

Hence derivatives of α are simply related to the derivatives of R_c as derived in Appendices 5.5.4 and 5.5.5.

$$\left(\frac{\partial \alpha}{\partial z_c}\right) = \frac{1}{B_c}\left(\frac{\partial R_c}{\partial z_c}\right) \tag{5.96}$$

$$\left(\frac{\partial \alpha}{\partial B_c}\right) = \frac{1}{B_c}\left(\frac{\partial R_c}{\partial B_c}\right) - \frac{R_c}{B_c^2} \tag{5.97}$$

$$\left(\frac{\partial \alpha}{\partial k_c}\right) = \frac{1}{B_c}\left(\frac{\partial R_c}{\partial k_c}\right) \tag{5.98}$$

5.5.7 Thick Clouds

Ideally, the retrieval would be able to converge to the same solution for any set of initial guess parameters. However, this is not true in this case because of the highly non-linear dependence of radiance on cloud optical thickness.

In the optically thick limit ($k_c \to \infty$, although in practice, $k_c \simeq 0.1$/km is sufficient) the pencil beam transmittance $\tau \to 0$ (Eq. 5.57), as do its derivatives w.r.t. macroscopic parameters cloud top height z_c and extinction k_c.

This means that the derivatives of the pencil beam radiances dL/dz_c (Eq. 5.78) and dL/dk_c (Eq. 5.79) also tend to zero, although when including the FOV convolution, some dependence on cloud top height is recovered due to the additional term $L_j(\partial\phi_j/\partial z_c)$ in Eq. (5.94).

Thus, when starting with an initial guess assuming thick cloud, it is difficult to iterate to retrieve any other value of cloud extinction. The retrieval is effectively reduced to just a two-parameter state vector of cloud top height z_c and cloud top radiance B_c.

6 Micro Retrieval

6.1 Determination of cloud particle size

On basis of the CSDB two methods have been assessed to derive information on particle size of cirrus clouds and PSCs. These methods are based on the observation (see TN [1], Figures 83-91) that for some brightness temperature difference correlations there appears a separation of particles with different effective radii.

Figures 6.1, 6.2 and 6.3 show typical examples for the correlations between size and radius for cirrus, ice PSCs and NAT PSCs. In case of cirrus a linear regression can be fitted to particle radii smaller than about 40 μm. Thus for the smaller radii a distinction seems feasible, however, with relatively large uncertainties as indicated by the 1-σ standard deviation of the datapoints. For larger cirrus particles a size discrimination is not possible. In case of PSCs the smaller variability and good correlations also indicate the possibility for a size determination.

In the following, two methods for retrieval of particle size are proposed. Both methods are based on normalised equivalent blackbody temperature-differences (ΔBT^{norm}) calculated from three different microwindows:

$$\Delta BT^{norm} = \frac{BT_1 - BT_2}{BT_3}. \tag{6.1}$$

Here BT_i is the equivalent blackbody temperature for microwindow i determined as:

$$BT_i = BT\left(\frac{1}{N_i}\sum_{j=1}^{N_i} S(\nu_{i,j})\right) \tag{6.2}$$

where N_i is the number of spectral grid points in microwindow i and $S(\nu_{i,j})$ is the radiance (of the apodized spectrum) at wavenumber $\nu_{i,j}$. Note that one microwindow does not necessarily consist of consecutive spectral grid points but may consist of various sub-microwindows.

6.1.1 Linear regression

A linear fit to the simulated normalised equivalent blackbody temperature-differences from the database is performed to get the input quantities a_{itype} and b_{itype} to be used within the processor.

$$R_{eff} = a_{itype} + b_{itype} \times \Delta BT^{norm}_{itype} \tag{6.3}$$

where $itype$ indicates the cloud type: cirrus, ice-PSC or NAT-PSC.

The values of a_{itype} and b_{itype} are stored as a pre-calculated look-up table and are an input to the processor.

6.1.2 Quadratic fit

This scheme uses a least-squares fitting of measured equivalent blackbody temperatures BT_i^{meas} in distinct microwindows i (total number: N_i) to mean equivalent blackbody temperatures derived from the database binned for different ranges (δci) of cloud-index values and for different effective radii:

$$BT_i^{CSDB}(\delta ci, R_{eff}) = \frac{1}{N_{\delta ci}} \sum_{k=1}^{N_{\delta ci}} BT \left(\frac{1}{N_i} \sum_{j=1}^{N_i} S_{k,R_{eff}}^{CSDB}(\nu_{i,j}) \right). \tag{6.4}$$

The values of $BT_i^{CSDB}(\delta ci, R_{eff})$ are stored as a pre-calculated look-up table separately for cirrus, ice-PSCs and NAT-PSCs and are an input to the processor.

The retrieval procedure calculates, for each effective radius of the look-up table within the actual CI-bin δci_{act} the sum of the mean square value of the differences:

$$\Delta^2(R_{eff}) = \sum_{i=1}^{N_i} \left(BT_i^{meas} - f_{min} \times BT_i^{CSDB}(\delta ci_{act}, R_{eff}) \right)^2 \tag{6.5}$$

where f_{min} is the wavelength-independent fit parameter determined as:

$$f_{min} = 1 + \frac{\sum_{i=1}^{N_i} \left[BT_i^{CSDB}(\delta ci_{act}, R_{eff}) \times \left(BT_i^{meas} - BT_i^{CSDB}(\delta ci_{act}, R_{eff}) \right) \right]}{\sum_{i=1}^{N_i} BT_i^{CSDB}(\delta ci_{act}, R_{eff})}. \tag{6.6}$$

The effective radius belonging to the smallest value of $\Delta^2(R_{eff})$ is the result of the procedure.

6.1.3 Restrictions

Both methods described are applied only to selected cloud cases with the following restrictions:

- Cirrus cloud, ice-PSC, or NAT-PSC
- Cloud-index-interval $ci_{min} < ci < ci_{max}$
- Cloud-height relative to CTH of the limb-scan.
  Tangent altitude index: $ihtang(CTH) - nhtang \geq ihtang \geq ihtang(CTH)$
- Absolute tangent altitude: $htang \geq h_{topopause} - \Delta_{tropopause}$

ci_{min},ci_{max},$nhtang$, $\Delta_{tropopause}$ are input-parameters to the processor, while ci, $ihtang(CTH)$, and $h_{topopause}$ are determined internally.

6.1.4 Internal validation

Figures 6.4, 6.5 and 6.6 show example results of the linear regression and the quadratic fit retrieval applied to the CSDB spectra. The cases have been performed for a cloud index range of 1.8-5.0 and restricted to the top two tangent heights affected by clouds. In general the linear regression method results in a good estimation of the mean value of the ensamble under investigation, however, with a large variability of the results. The accuracy of the quadratic fit method depends largely on the number of microwindows included. In case of cirrus (Fig. 6.4) the spectral regions around 827 and 941 cm^{-1} lead to the best results while for ice- and NAT-PSC 827, 941 and 1227 cm^{-1} are favourable. In general the variability of the results of the quadratic fit method is smaller as can be seen from the error bars in the right columns of Figs. 6.4, 6.5 and 6.6. Additionally, when applied to different ci-ranges and tangent altitude restrictions (not shown here), the results are in general more stable.

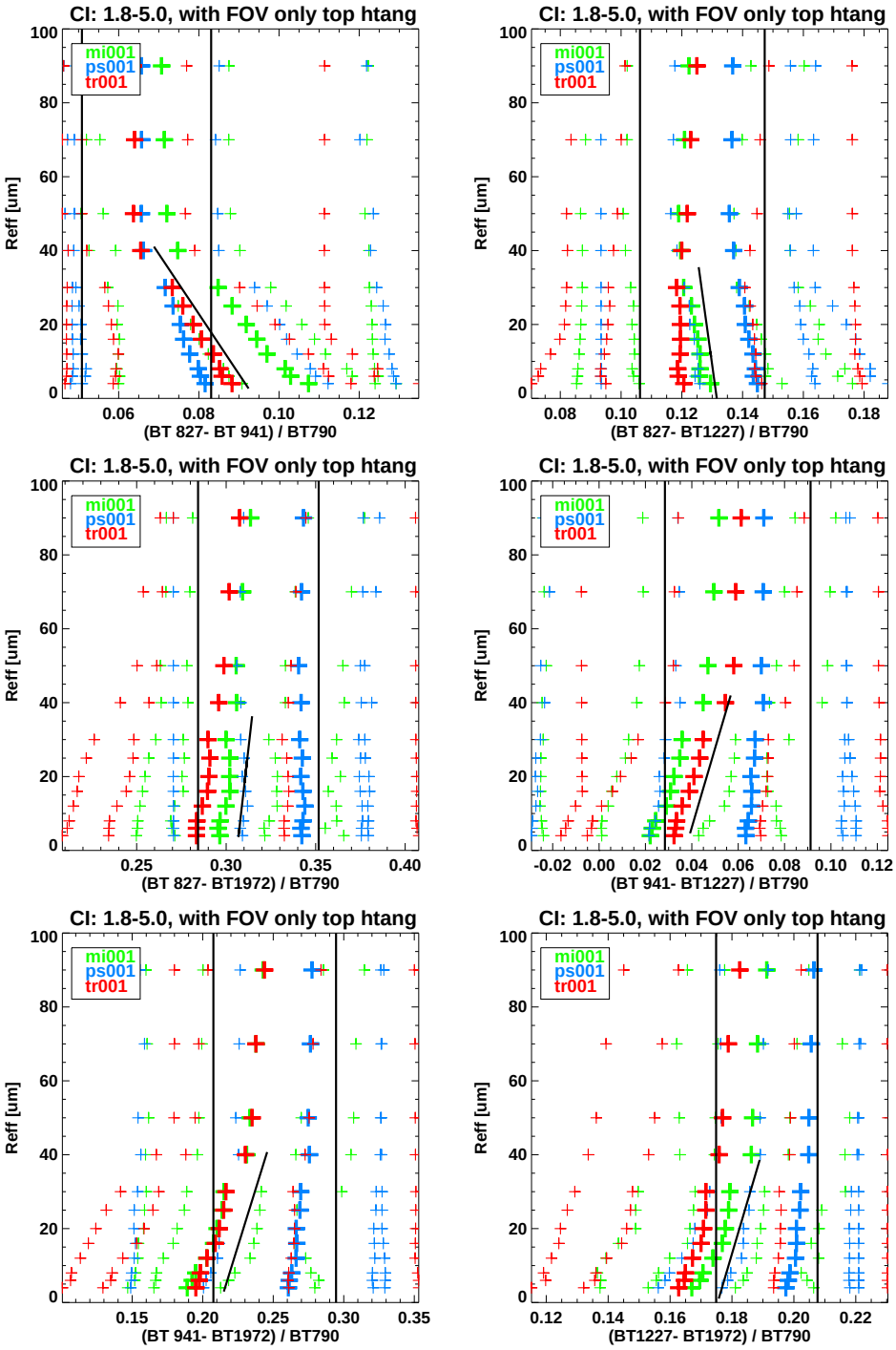

Figure 6.1: Correlation between normalized equivalent blackbody temperature differences for different wavenumber regions and the effective radius in case of cirrus clouds. Different colors denote different atmospheres (mi001=mid-latitudes, ps001=polar summer, tr001=tropics). The bold crosses are the mean values while the smaller symbols right and left of the bold ones denote the 1-σ variability and the max/min values. Slant black lines are fits to the data for effective radii $\leq 40\,\mu m$ and vertical black lines indicate the variability of the mean above 40 km.

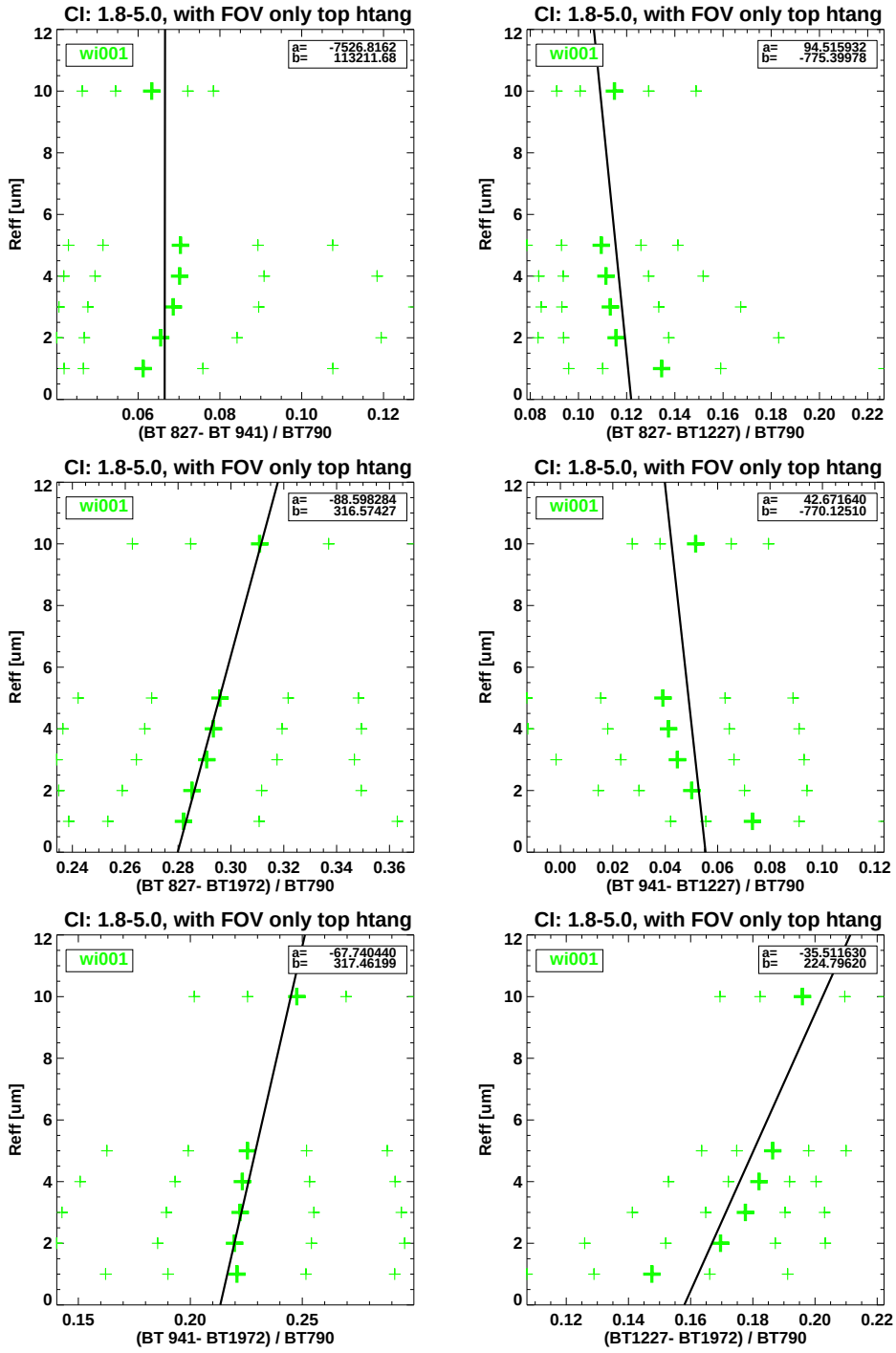

Figure 6.2: Like Fig. 6.1, but for ice-PSCs and one atmosphere (polar winter).

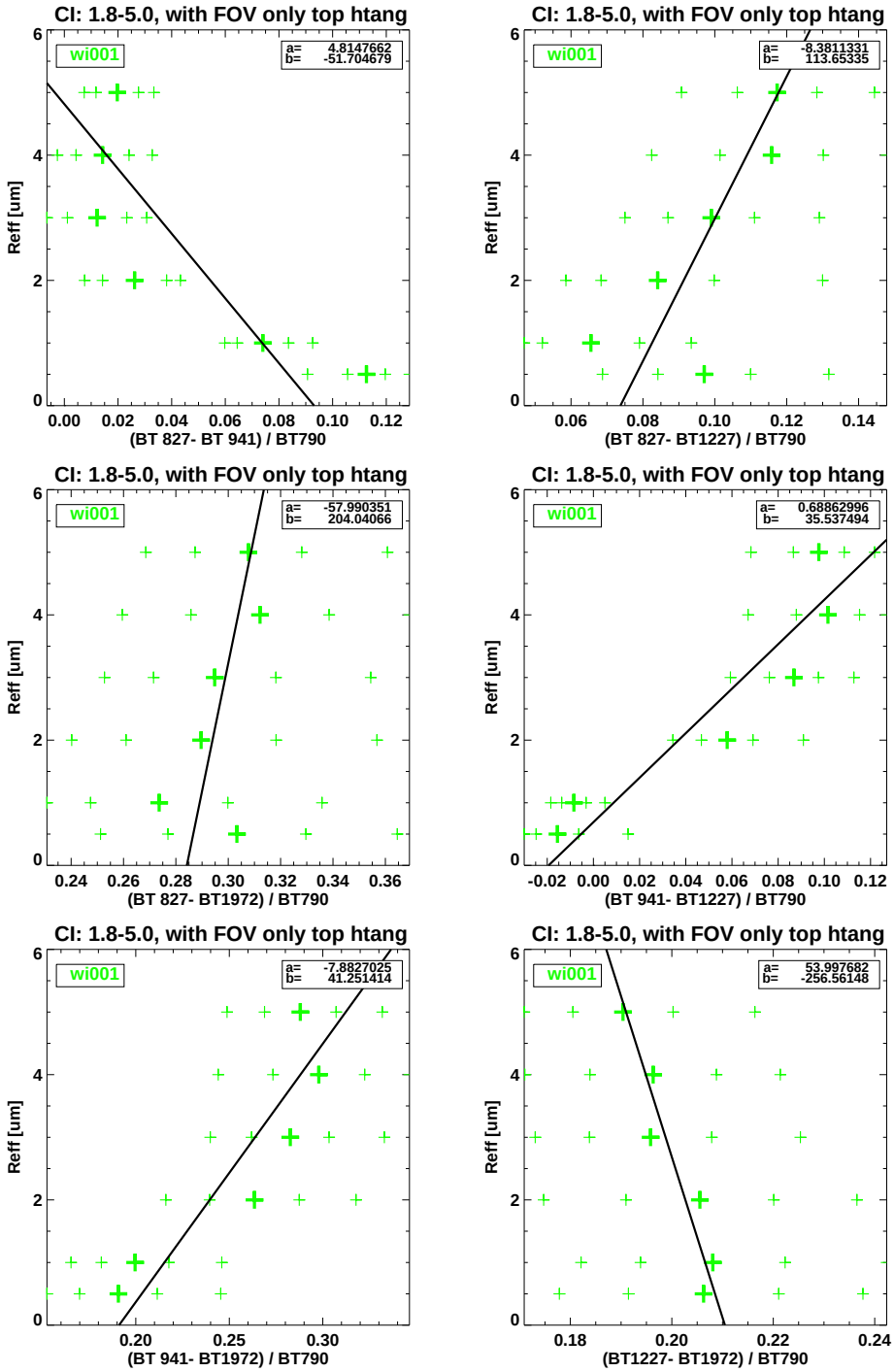

Figure 6.3: Like Fig. 6.1, but for NAT-PSCs and one atmosphere (polar winter).

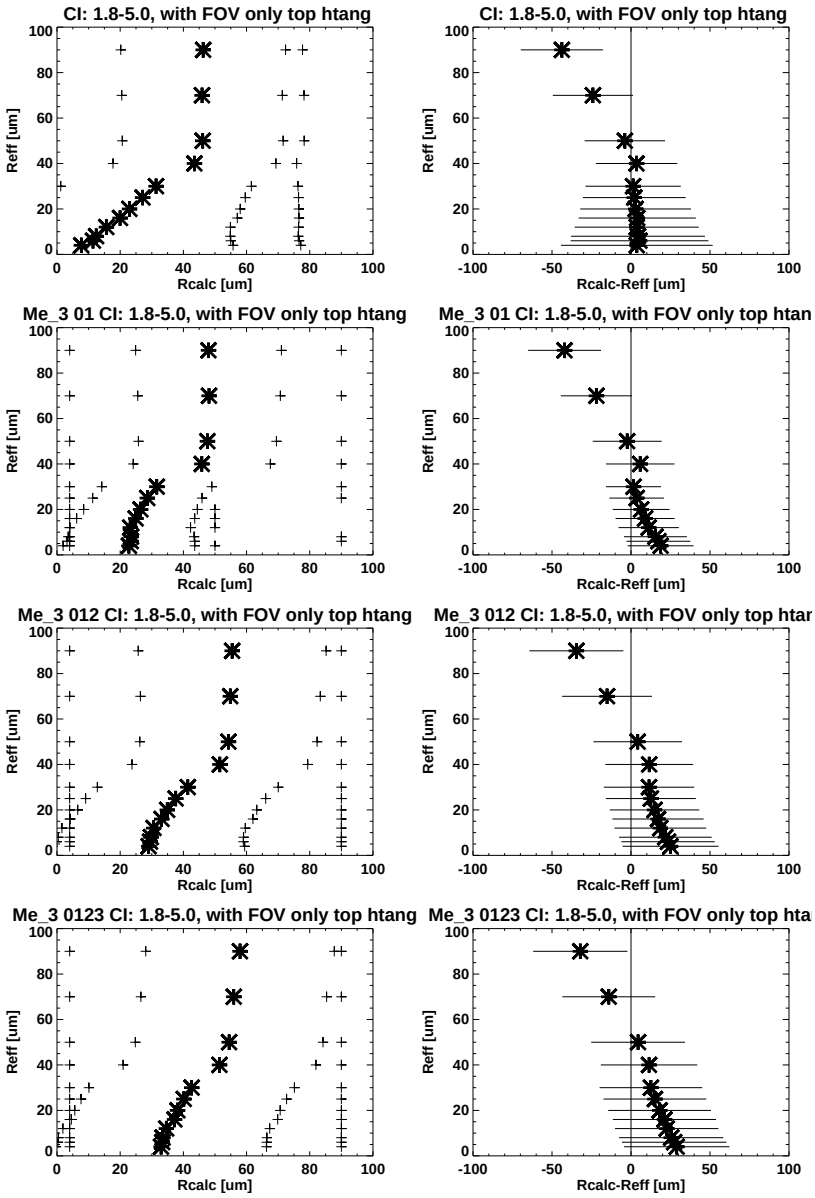

Figure 6.4: Comparison of derived effective radii with exact values from the CSDB. Left: correlation between calculated radius (Rcalc) and correct value (Reff). Right: difference between calculated and correct value: crosses are the mean differences while the error bars indicate the 1-σ variability of the differences. Rows: top: linear regression method based on the normalized equivalent blackbody temperature differences (BT827-BT941)/BT790; 2nd-bottom row: quadratic fit on basis of microwindows around 827 and 941 cm^{-1} (2nd row), 827, 941 and 1227 cm^{-1} (3rd row) and 827, 941, 1227, and 1971 cm^{-1} (bottom row).

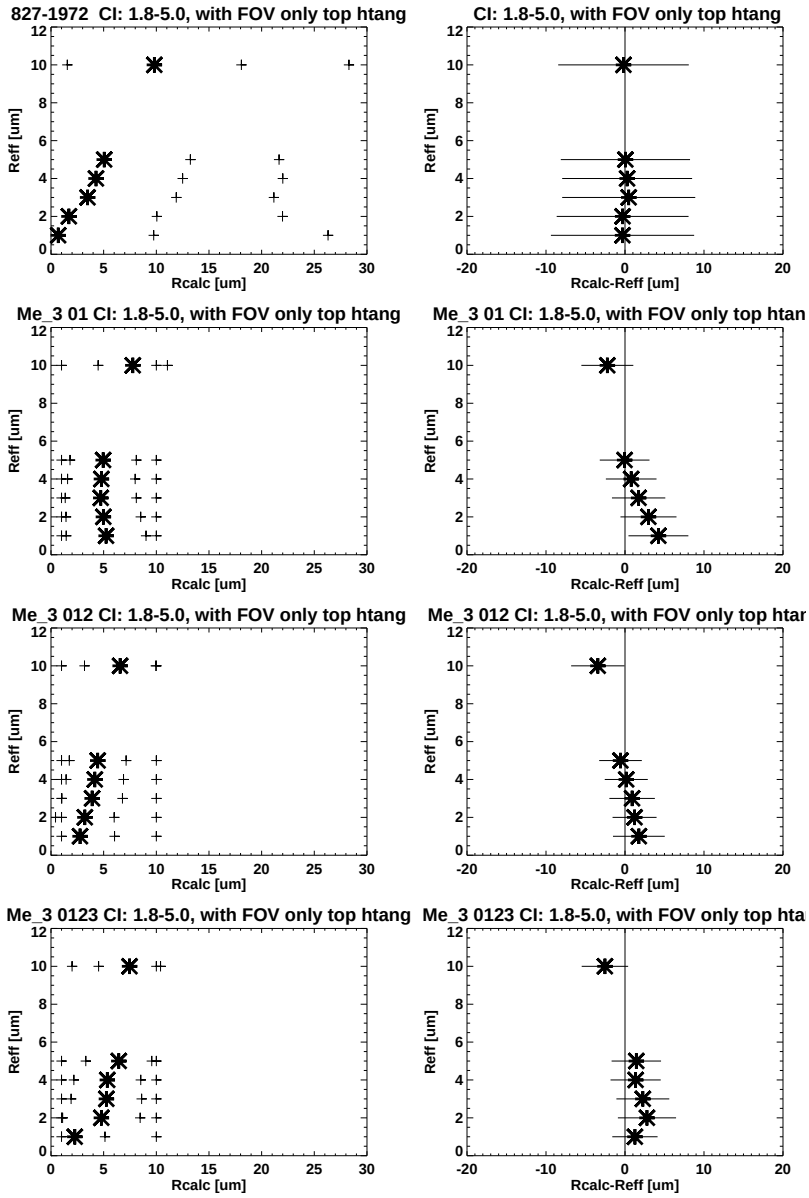

Figure 6.5: Like Fig. 6.4 but for ice-PSCs. For the first row the linear regression has been performed for the (BT827-BT1972)/BT790 differences.

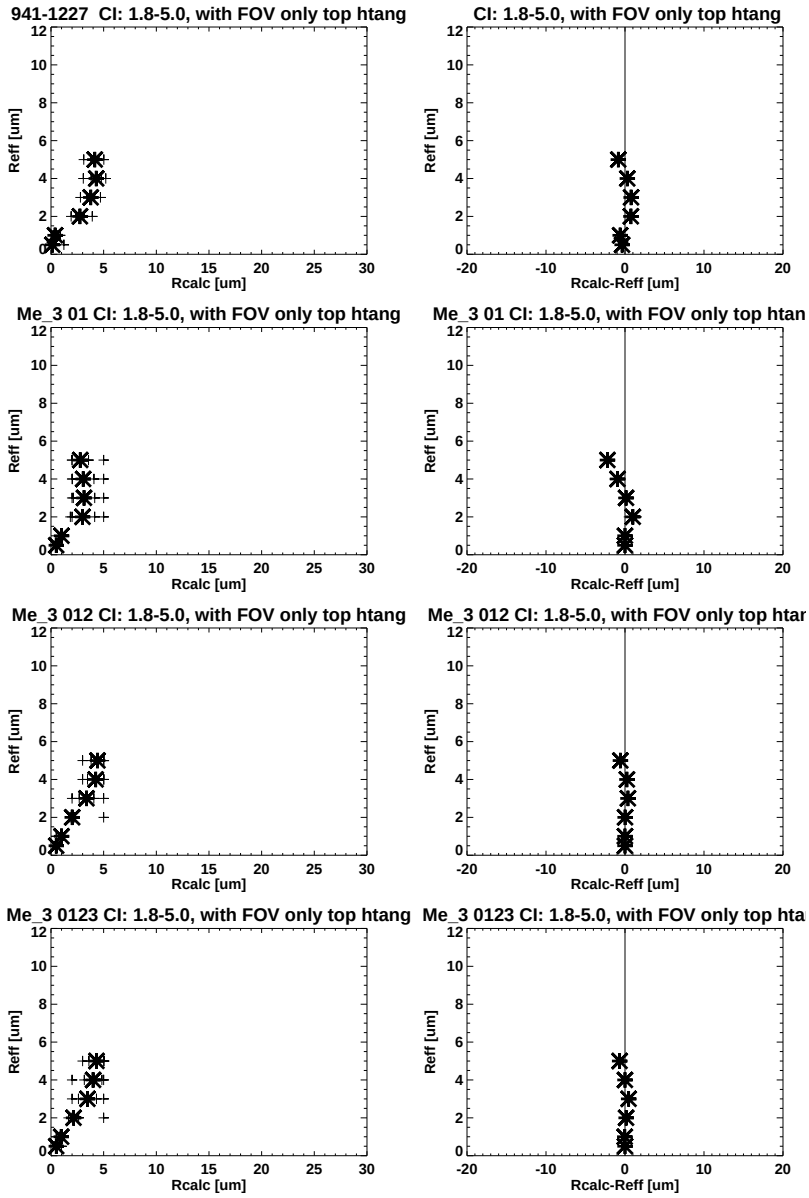

Figure 6.6: Like Fig. 6.4 but for NAT-PSCs. For the first row the linear regression has been performed for the (BT941-BT1227)/BT790 differences.

6.2 Estimates of Volume and Area Density Quantities

The absorption and extinction characteristic of a cloud is mainly dominated by the surface area density (AD, A) integrated along optically path. Analyses with the CSDB show, that CI_A is very well correlated with the limb integrated are density, called here area density path (ADP). This quantity, is under the assumption, that the cloud is filling the complete FOV, directly linked to radiance measurements and the sensitivity of an instrument to measure clouds up to a certain ADP, before a cloud is optically thick and higher densities are not able to differentiate.

Area density (A) is defined as:

$$A = \frac{3 \cdot V}{R_{eff}}$$

with V: Volume density typically in $[\mu m^3/cm^3]$ and R_{eff}: the effective radius in $[\mu m]$ of the PSD, and A typically in $[\mu m^2/cm^3]$. The relation is exactly correct only for spherical particles. The quantities necessary to calculate A for the modelled spectra have been defined in the CSDB.

ADP is simply the summation of the products of a single limb path segment length times the AD for the respective altitude, or the integrated AD from the observer to the tangent point and to deep space:

$$ADP = \int_{obs=0}^{\infty} A\,dx \qquad [\mu m^2/cm^2]$$

The uncertainty in the observer geometry for the measurement like

1. where is the cloud located along the limb path?
2. how large is the horizontal extent of the cloud?

is eliminated by this approach. Therefore, ADP results in an excellent quantity for comparisons with global models like ECWMF, where cloud physics parametrisation are incorporated and the limb path can be traced through the 3D model output of IWC or AD.

If radius information can be retrieved $R_{eff,ret}$, like for example by the method described in Section 6.1, Volume densities (or the equivalent IWC or LWC) are straight forward to derive with the formula above. With the the simplification of a 1D-atmosphere, the cloud layer is assumed to fill the complete tangent height (TH) layer. With the path length in the TH layer L_{TH} for a certain FOV width FOV (e.g. 4 km) :

$$L_{TH} \cong \sqrt{4 \cdot FOV \cdot (r_e + TH)}\ ,$$

with r_e: earth radius, it is possible to estimate the area and volume density:

$$A = \frac{ADP}{L_{TH}}$$

$$V = \frac{ADP \cdot R_{eff,ret}}{3 \cdot L_{TH}}$$

The excellent correlation between CI-A and ADP permits a simple approach for an estimate of ADP in the measurements. Like illustrated in Figure 6.7 ADP shows a nearly linear relation with CI-A in the CI-range 1.15 to 5. A polynomial fit will improve the alignment with the scattered data points. The scatter band along the

fit is a good estimate for the uncertainty of the method, which is typically in the order of $\pm$50%. The relation is saturated above a certain threshold value for CI-A, which is a quite stable value for different conditions (in the Fig. 6.7a, $CI_A < 1.2 \rightarrow ADP_{thres} \cong 10^8\ \mu m^2/cm^2$). Obviously from Figure 6.7a and b the fit parameter and threshold value are quite constant with altitude an diffrent background atmospheres, if the line of sight passes through a TH layer filled with a cloud. When the cloud layer is signifcant above the TH then the paramter will change or the scater band is getting broader. Optimised threshold values and fit parameter will be prepared for look-up tables.

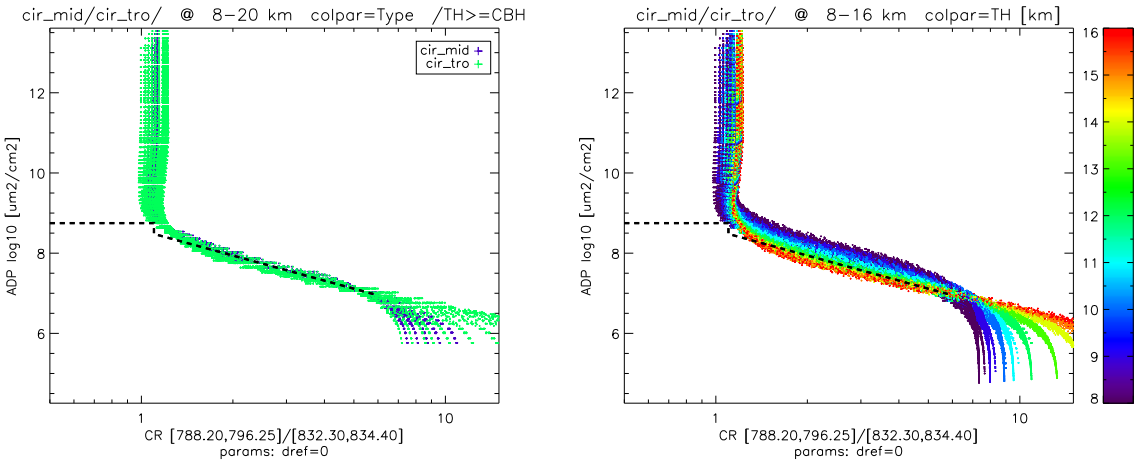

Figure 6.7: CI-A versus Area Density Path (ADP) for (left) mid-latitude and equatorial CSDB spectra in the altitude range 8-20 km (Only $TH \geq CBH$ is considered). (right) In contrast with no restriction for CBH but from 8-16 km with colour-coding for the tangent heights.

The whole approach presumes that the instrument FOV is completely filled with cloud. Any partially coverage of the FOV results in a mixture of cloud-free and cloudy radiances, which changes the CI-A and a realistic ADP is difficult to quantify. The best possible approach is to use the CTH_{macro} information form the macroscopic retrieval (Section 5) for a pre-selection of spectra with low cloud extent in the FOV. CTH_{macro} takes into account the cloud effective fraction in the FOV. The analysis should be restricted to the same selected cloud cases like in the R_{eff} retrieval (section 6.1.3) but can be applied to classified liquid cloud spectra as well.

7 Error and Quality Assessment

For most of the algorithms described above it is not possible to retrieve standard error quantities (like covariance matrices or error propagation of systematic and random error sources) due to the fact that the parameter has no real physical error (e.g. cloud types or the detection of clouds). For these parameters a measure of confidence has been defined. Therewith data users are able to evaluate the quality and uncertainty of the retrieval parameters.

7.1 Cloud Detection

A qualitative measure of the confidence for the detection of a single cloud event at a specific TH is generated by the different cloud detection methods, where each method results in a cloud flag $\mathbf{FLAG}_{method}$ with value $cloudy = 1$ or $non-cloudy = 0$. In addition the information flags like the scattering index and water vapour index (for details see chapter 3) or the temperature criteria flag (section 2.4.4) are used to enhance or reduce the level of confidence, where flags reducing the confidence have to be considered by substraction (or by definition of a negative sign for this flag). The total of all flags $CONF$ (as described in section 3.11) gives a measure of confidence:

$$CONF = \sum_{i=0}^{N_{flag}} \mathbf{FLAG}_i$$

Optically very thin clouds will result in smaller confidence values. For the recent number of five detection methods a maximum of $CONF = 5$ highlights the maximum confidence level. Negative values are possible under certain conditions (e.g. only one methods detects a cloud but temperatures are significant to high for cloud formation and the water vapour emissions are high as well $\rightarrow CONF = -1$). The normalized confidence level as described in section 3.12 gives a relative confidence for each cloud event at a specific TH depending on the number of detection methods.

7.2 Classification

7.2.1 Flags from Scatter Diagrams

Main information of the cloud classification will come from the Bayes analysis, where error quantities for the training dataset (CSDB) can be retrieved (see next section). Additional information from the simple two parameter scatter diagrams (e.g CI vs BTD), the cloud type flags, will add on confidence in the interpretation of the spectra. Similar to the cloud detection confidence levels for the classification can be generated by simple summation of the cloud type flags (see sections 4.2 and 4.3).

7.2.2 Bayes Histograms

7.2.2.1 Accuracy of Bayes Classification

The error assessment of a primitive Bayesian classification for a set of possible outcomes $\{C_1, C_2, \ldots, C_d\}$ is measured by the quantities of the *confusion matrix* [23]. The matrix provides an estimation of the classification quality. The results of the comparison along the classification result and reference data for each class is collected. The confusion matrix contains the classification results (rows) and the reference data (columns) where n_{ij} denotes the number of objects of class C_j classified as part of class C_i. Consequently the diagonal entries specify a correct classification. Table 7.1 shows a schematic example.

Classification \ Reference	C_1	$\cdots$	C_d	$\sum$ classification
C_1	n_{11}	$\cdots$	n_{1d}	$\sum_{j=1}^{d} n_{1j}$
$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
C_d	n_{d1}	$\cdots$	n_{dd}	$\sum_{j=1}^{d} n_{dj}$
$\sum$ reference	$\sum_{i=1}^{d} n_{i1}$	$\cdots$	$\sum_{i=1}^{d} n_{id}$	$\sum_{i,j=1}^{d} n_{ij}$

Table 7.1: confusion matrix

The most simple value that can be derived is the ***overall accuracy***, which measures the correct classified values to the total amount of analyzed values. To get a percentage the formula ends in

$$oaa = \frac{\sum_{i=1}^{d} n_{ii}}{\sum_{i,j=1}^{d} n_{ij}} \cdot 100$$

To separately estimate the classification accuracy for each class C_k the *producer's* and *user's accuracy* may be measured.

The ***producer's accuracy*** describes the part of reference data which is covered in the classification

$$pa_k = \frac{n_{kk}}{\sum_{i=1}^{d} n_{ik}} \cdot 100 = \frac{\text{amount of correctly classified values for } C_k}{\text{total amount of reference data values in } C_k}$$

The ***user's accuracy*** describes the percentage of the captured data that fit the classification .

$$ua_k = \frac{n_{kk}}{\sum_{j=1}^{d} n_{kj}} \cdot 100 = \frac{\text{amount of correctly classified values for } C_k}{\text{total amount of classified reference data values in } C_k}$$

This two accuracies cover the error of commission, the wrongly classified elements, as well as the error omission, the not classified elements of a class. Note that both accuracies have the same value if all objects of a class could be classified, not saying that the classification was correct.

An additional degree of conformance is the ***Cohen's kappa coefficient*** which denotes the intensity of correlation between the classification result and the reference data. The coefficient ranges from -1 to 1, where -1 accounts for a completely random consensus and 1 excludes a random consensus between reference data and classification results. The formula is given by the ratio

$$\kappa = \frac{\sum_{i,j=1}^{d} n_{ij} \cdot \sum_{k=1}^{d} n_{kk} - \sum_{k=1}^{d} (n_{\cdot k} \cdot n_{k\cdot})}{(\sum_{i,j=1}^{d} n_{ij})^2 - \sum_{k=1}^{d} (n_{\cdot k} \cdot n_{k\cdot})} = \frac{\text{reached accuracy} - \text{random assignment}}{1 - \text{random assignment}}$$

with $n_{\cdot k} = \sum_{i=1}^{d} n_{ik}$ and $n_{k\cdot} = \sum_{j=1}^{d} n_{kj}$

The next tables show the different accuracies for the examples of the bayes analysis given in chapter 4.2.3 and 4.3.2. Table 4.3 was an example of a Bayes test analysis for ice-STS-NAT with all CSDB spectra and only for optically thick conditions ($CI_A < 1.2$) for histograms with 4 K bins. Table 7.2 lists the different accuracy parameters for the three classes in both cases achieved while training the classifier.

Accuracy / Class	OAA[%]	UA[%]	PA[%]	κ
	94.97			1
all ice		99.61	99.61	
all STS		96.46	96.46	
all NAT		90.01	90.01	
	100			1
thick ice		100	100	
thick STS		100	100	
thick NAT		100	100	

Table 7.2: Accuracies for Bayes test analysis for ice-STS-NAT with all CSDB spectra and only for optically thick conditions ($CI_A < 1.2$) for histograms with 4 K bins.

For the case where all spectra have been used for training the classification of the data trained with reaches a relatively high overall accuracy of nearly 95% and a clear non random consensus between reference data and classification results. The user's and producer's accuracies show the same grading as the classification results in Table 4.3. Same interpretation holds for the optically thick case where the results even show 100 percentage values.

Table 4.4 was an example of Bayes analyses for liquid and cirrus spectra of the CSDB. Training proceeded with all, optically thin events ($CI_A < 1.8$), thin ($1.8 < CI_A < 5$) scenarios for mid-latitude and tangent heights greater than the cloud base height (CBH), and the same extraction for polar summer scenarios for classification histograms with 4 K bins. Table 7.3 lists the different accuracy parameters for the two classes in all cases achieved while training the classifier.

Class \ Accuracy	OAA[%]	UA[%]	PA[%]	κ
	75.42			0.505
all cirrus		74.35	74.35	
all liquid		76.69	76.69	
	97.36			0.947
thin cirrus		97.49	97.49	
thin liquid		97.21	97.21	
	98.75			0.975
thin/midl/CBH cirrus		97.87	97.87	
thin/midl/CBH liquid		99.55	99.55	
	96.09			0.919
thin/psum/CBH cirrus		98.2	98.2	
thin/psum/CBH liquid		93.08	93.08	

Table 7.3: Accuracies for Bayes test analyses for liquid and cirrus spectra of the CSDB. For training histograms optically thin events ($CI_A < 1.8$), thin ($1.8 < CI_A < 5$) scenarios for mid-latitude and tangent heights greater than the cloud base height (CBH), and the same extraction for polar summer scenarios. All analyses for classification histograms with 4 K bins.

The interpretation of the accuracies again leads to the classification results shown in Table 4.4. The overall accuracy in the case where all spectra are used for training and classification yields a value of $\approx 75\%$ which implies, that 3 out of 4 spectra are classified correctly. Both classes show a similar user's and producer's accuracy. Looking at the classification results the restriction of certain parameters improves the results.

An optimised set of classes will be prepared for the processor together with the corresponding statistical accuracy quantities described above.

7.2.2.2 Quality of the CSDB-based Classification Scheme

Because the data set used for the training of the classifier does not reflect the true distribution of the atmosphere (e.g. the microphysical parameter are spread equally in the parameter space but are not weighted with their occurrence in the atmosphere) a statistical approach has been applied to the CSDB spectra to validate the quality of the classification scheme and is presented below.

A set of data which is taken for training of the classifier and creation of the histograms for a class may be divided into two parts randomly (e.g. 10 and 90% of the spectra). With a subset containing 90% of the data the classifier is trained and the optimal ratios are chosen. The other 10 percent are known to belong to this class but are not used to train the classifier. Hence, this data is taken as the input for an independent classification test. The result of this classification gives information how unerring the method is for a kind of *blind test* data set of spectra (not used for the training of the classification). A statistical error quantity for a classification for a specific cloud type j based on the full dataset p_{100_j} (e.g. Table 4.3) can be given by doing the selection analysis N times and then compute the mean absolute difference of these results i in respect to the value of the analysis with the complete dataset.

$$\Delta p_{100_j} = \frac{\sum_{i=0}^{N} \mid p_{100_j} - p_{90_{i,j}} \mid}{N}$$

Table 7.4 shows the results for the analysis of PSC data sets that were not employed for the training. With $N = 10$ the table shows the error with respect to the results of table 4.3.

ice/STS/NAT	Ice	NAT	sa0248	sa2525	sa4802
all	ia: 0.21 na: 0.2 sts: 0.02	ia: 0.00 na: 0.53 sts: 0.53	ia: 0.27 na: 1.01 sts: 0.72	ia: 0.38 na: 2.57 sts: 3.35	ia: 0.15 na: 0.45 sts: 2.62
thick	ia: 0.00 na: 0.00 sts: 0.00	ia: 0.00 na: 0.00 sts: 0.00	ia: 0.17 na: 0.00 sts: 0.17	ia: 0.00 na: 0.96 sts: 3.32	ia: 0.00 na: 4.48 sts: 5.11

Table 7.4: Error example of Bayes test analysis for ice-STS-NAT with CSDB spectra. error in percent (ice: ia, NAT: na and various STS: sa compositions) for histograms with 4 K bins.

The error measure can be interpreted as follows: For example the training with all spectra and analysis with all spectra yields that, 99.61% could correctly be classified as ice, whereas 0.39% were classified as NAT and no spec was classified as sts. The error accounts for a certainty that independent ice spectra can be classified as ice with $99.61\% \pm 0.21\%$, classified as NAT with $0.39\% \pm 0.2\%$ and classified as STS with 0%.

Table 7.5 shows the results for the 10% of cirrus and liquid spectra data sets that were not employed for the training. With $N = 10$ the table shows the error with respect to the results of table 4.4. Obviously the confidence in classification is significantly reduced for the restricted training datasets (row three and four). Value greater than 10 indicates large uncertainties in the corresponding retrieved probability p_{100_j}.

cirrus/liquid	cir-mid	cir-psum	cir-tro	liq-mid	liq-psum	liq-tro
all	cir: 2.23 liq: 4.37	cir: 2.37 liq: 2.37	cir: 0.73 liq: 3.74	cir: 4.01 liq: 4.88	cir: 10.93 liq:10.88	cir: 1.07 liq: 3.96
thin	cir: 10.65 liq: 20.35	cir: 31.03 liq: 17.17	cir: 0.79 liq: 1.71	cir: 4.03 liq: 23.85	cir: 25.95 liq: 14.05	cir: 1.52 liq: 11.87
thin/midl/ CBH	cir: 1.14 liq: 1.13	cir: 4.23 liq: 10.25	cir: 19.32 liq: 17.61	cir: 1.04 liq: 3.78	cir: 19.28 liq: 17.35	cir: 2.65 liq: 6.41
thin/psum/ CBH	cir: 3.04 liq: 5.86	cir: 0.18 liq: 0.37	cir: 4.17 liq: 3.92	cir: 3.31 liq: 1.88	cir: 2.51 liq: 2.79	cir: 3.11 liq: 0.91

Table 7.5: Error example of Bayes test analyses for liquid and cirrus for histograms with 4 K bins in respect of Table 4.4 .

7.3 Macroscopic Retrieval Parameter

The Macropscopic Parameter retrievals are carried out for a number ($\sim$5) of microwindows, each generating its own set of three retrieved parameters (CTH, CEX, CTT) $\mathbf{x}_i$ and associated covariance matrices $\mathbf{S}_i$

Ignoring any that were flagged as unsuccessful, a weighted average $\hat{\mathbf{x}}$ and covariance $\hat{\mathbf{S}}$ is constructed from

the remaining retrievals

$$\hat{\mathbf{x}} = \sum_i \mathbf{S}_i{}^{-1}\mathbf{x}_i \tag{7.1}$$

$$\hat{\mathbf{S}}^{-1} = \sum_i \mathbf{S}_i{}^{-1} \tag{7.2}$$

A spike test is then be performed on each of the input retrievals

$$\chi_i^2 = (\mathbf{x}_i - \hat{\mathbf{x}})^T \mathbf{S}_i{}^{-1} (\mathbf{x}_i - \hat{\mathbf{x}}) \tag{7.3}$$

If the maximum value of χ_i^2 exceeds some predefined value (e.g., 9, corresponding to a '3-sigma spike test' in variance rather than SD) then that retrieval is also removed and the process repeated from Eq. (7.1). This continues until either all remaining results are consistent within the boundaries of the spike test, or even the remaining two results are inconsistent to this extent.

The final stage is to 'inflate' the combined covariance $\hat{\mathbf{S}}$ so that the diagonal elements match the variances of the remaining results. This is done by constructing a vector of the ratios **r**

$$r_j = \sqrt{\frac{\sigma_j^2}{\hat{S_{jj}}}} \tag{7.4}$$

where σ_j^2 is the variance of the retrieved parameter j in the remaining results, and $\hat{S_{jj}}$ the corresponding diagonal element of $\hat{\mathbf{S}}$. The elements r_j will be limited to a minimum value 1 (since it is possible for two retrievals to agree, coincidentally, within their theoretical uncertainties).

The adjusted covariance of the macroscopic retrieval is then given by

$$\hat{\mathbf{S}}' = \mathbf{r}\,\hat{\mathbf{S}}\,\mathbf{r}^T \tag{7.5}$$

This will then provide an estimate of the uncertainties in each of the retrieved parameter (CTH,CTT,CEX) consistent with the observed scatter in results from the different microwindows after the obvious outliers have been removed. Due to the highly correlated nature of the macroscopic parameter retrieval, it will be important to present the full matrix rather than just variances of each parameter.

A further (integer) retrieval quality flag will be provided representing the number of retrievals that were flagged as bad or rejected by the spike test (0=best, 5=worst, assuming 5 microwindows used and all failing the convergence test).

7.4 Micro Retrieval Parameter

7.4.1 Radius Retrieval

For the error assessment of the parameterised cirrus size retrieval a new test dataset comprising only the used spectral windows will be compiled. Since this test-CSDB will only contain small microwindows and will

not include any Jacobian calculations it can cover a much larger space of atmospheric variability (temperature, humidity) than the broadband CSDB itself.

The error assessment of the size retrieval will then be performed on the evaluation of selected random samples of this larger test-dataset. In case of large deficiencies, the parameterisation of the radius retrieval will be improved on basis of the original CSDB and the new test-CSDB.

A further part to assess the errors is the comparison with the validation dataset of simulations with realistic cloud fields including multiple scattering as done in the feasibility study [1]. Due to restrictions of the calculation time, this validation dataset cannot be as large as the test-CSDB and, thus, the statistics is not as large as for the test-CSDB. Thus, this approach will result in errors of the retrieval approach due to non-homogeneous clouds and multiple scattering effects while the test-CSDB error estimation comprises the inherent errors of the retrieval approach.

7.4.2 Area Density Path Retrieval

Errors for ADP can be best estimated with the same approach like for the radius retrieval. Therefore the extended test dataset will also include the MWs necessary for cloud index calculations. However, a first estimate of the ADP-method error is the root mean square error in respect to the retrieved fit-curve adp of $\log(CI)$ versus $\log(ADP)$, which is in the order of:

$$\Delta\log(adp) = RMSE \implies \Delta adp(CI) = \begin{cases} adp(CI) - 10^{\log(adp(CI)) - RMSE} \\ 10^{\log(adp(CI)) + RMSE} - adp(CI) \end{cases}$$

where the roots mean square error (RMSE) is given for each logarithmic fit in the specific altitude and latitude range. For example for tropical conditions, at 14 km altitude, and a middle CI=3, the error results in ΔADP of +11.9% or -13.5% respectively (see examples in Figure 7.1). The errors increase significantly for lower altitudes, especially in the tropics. Input files for the prototype processor with the fitting coefficients will include the RMSE values for easy computation of the ADP fitting error.

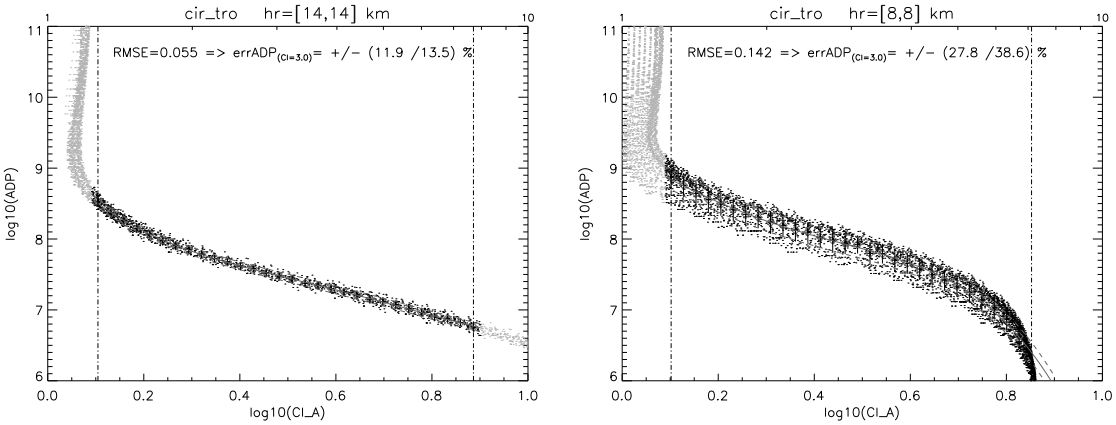

Figure 7.1: CI-A versus Area Density Path (ADP) equatorial CSDB spectra at 14 km and 8 km altitude. Dashed line represents the RMSE error. Symbols with error bars are representing the bins for the polynominal fit including the standard deviation. The RMSE value represents the $\log(ADP)$-value.

List of Figures

List of Tables

Bibliography

[1] Spang, R., Griessbach, S., Höpfner, H., Dudhai, A., Hurley, J., Siddans, R., Waterfall, A., Remedios, J.J., Sembhi, H.,
Technical Note: Retrievability of MIPAS cloud parameter
ESRIN Contract No.:20601/07/I-OL, 2008.

[2] Spang, R., Griessbach, S., Höpfner, H., Dudhai, A., Hurley, J., Siddans, R., Waterfall, A., Remedios, J.J., Sembhi, H.,
Product Validation Plan for the MIPAS cloud parameter processor, ESRIN Contract No.:20601/07/I-OL, 2008.

[3] Adriani, A., Massoli, P., Di Donfrancesco, G., Cairo, F., Moriconi, M. L., and Snels, M.
Climatology of polar stratospheric clouds based on lidar observations from 1993 to 2001 over McMurdo Station, Antarctica, J. Geophys. Res.
109(D24), D24 211, doi:10.1029/2004JD004800, 2004.

[4] Rossow, WB., RA Schiffer(1999)
Advances in understanding clouds from ISCCP
BULLETIN OF THE AMERICAN METEOROLOGICAL SOCIETY, 80, 2261-2287.

[5] Höpfner, M., N. Larsen, R. Spang, B. P. Luo, J. Ma, S. H. Svendsen, S. D. Eckermann, B. Knudsen, P. Massoli, F. Cairo, G. Stiller, T. v. Clarmann, and H. Fischer
MIPAS detects Antarctic stratospheric belt of NAT PSCs caused by mountain waves Atmos. Chem. Phys., 6, 1221-1230, 2006a.

[6] Höpfner, M., B. P. Luo, P. Massoli, F. Cairo, R. Spang, M. Snels, G. Di Donfrancesco, G. Stiller, T. von Clarmann, H. Fischer, and U. Biermann
Spectroscopic evidence for NAT, STS, and ice in MIPAS infrared limb emission measurements of polar stratospheric clouds
Atmos. Chem. Phys., 6, 1201-1219, 2006b.

[7] Remedios, J.
Profiles for MIPAS
EOS, Space Research Centre, Leicester, U.K., January 2001.

[8] Spang, R., M. Riese, G. Eidmann, D. Offermann, and P.H. Wang
A detection method for cirrus clouds using CRISTA 1 and 2 measurements
Adv. Space Res., 27/10, 1629- 1634, 2001.

[9] Spang, R., Remedios, J. J.
Observations of a distinctive infra-red spectral feature in the atmospheric spectra of polar stratospheric

clouds measured by the CRISTA instrument
Geophys. Res. Lett., 30(16), 1875, doi:10.1029/2003GL017231, 2003.

[10] Spang R. , J.J. Remedios, and M.P. Barkley
Colour indices for the detection and differentiation of cloud types in infra-red limb emission spectra, Adv. Space Res., 33 1041-1047, 2004.

[11] Spang R. , J. J. Remedios, L. J. Kramer, L. R. Poole, M. D. Fromm, M. Müller, G. Baumgarten, P. Konopka
Polar stratospheric cloud observations by MIPAS on ENVISAT: detection method, validation and analysis of the northern hemisphere winter 2002/2003
Atmos. Chem. Phys. 5, 679-692, 2005a.

[12] Spang, R., J. J. Remedios, S. Tilmes, and M. Riese
MIPAS observations of Polar Stratospheric Clouds in the Arctic 2002/3 and Antarctic 2003 winters
Advances in Space Research, 36, 868-878, 2005b.

[13] Dudhia, A.
The Reference Forward Model (RFM): Software User's Manual (SUM): `http://www-atm.physics.ox.ac.uk/RFM/sum.html`

[14] Greenhough, J., J. J. Remedios, H. Sembhi, and L. J. Kramer
Towards cloud detection and cloud frequency distributions from MIPAS infra-red observations
Advances in Space Research, 36, 800-806, 2005.

[15] Hurley, Jane
Modelling Clouds in the Infrared, thesis submitted in partial fulfillment of DPhil
University of Oxford, Oxford, UK, 2008.

[16] Press, W. et al.
Numerical Recipes: The Art of Scientific Computing
Cambridge University Press, **3**, pp.65–74, 2007

[17] Rodgers, C. D.
Inverse Methods for Atmospheric Sounding: Theory and Practice
World Scientific Publishing Co Pte Ltd, 2000.

[18] Stull, R. B.
Meteorology for Scientists and Engineers
Brooks Cole, **2**, pp.1-528, 2000.

[19] M. Kiefer and T. von Clarmann and U. Grabowski and M. De Laurentis and R. Mantovani and M. Milz and M. Ridolfi
Characterization of MIPAS elevation pointing
Atmos. Chem. Phys. 7, pp.1615-1628, 2007.

[20] T. von Clarmann and N. Glatthor and U. Grabowski and M. Höpfner and S. Kellmann and M. Kiefer and A. Linden and G. Mengistu Tsidu and M. Milz and T. Steck and G. P. Stiller and D. Y. Wang and H. Fischer and B. Funke and S. Gil-López and M. López-Puertas

Retrieval of temperature and tangent altitude pointing from limb emission spectra recorded from space by the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS)
J. Geophys. Res. 108 D23 4736, 10.1029/2003JD003602, 2003.

[21] M. Ridolfi and B. Carli and M. Carlotti and T. von Clarmann and B. Dinelli and A. Dudhia and J.-M. Flaud and M. Höpfner and P. E. Morris and P. Raspollini and G. Stiller and R. J. Wells
Optimized Forward and Retrieval Scheme for MIPAS Near-Real-Time Data Processing
Appl. Opt. 39, 8, pp.1323-1340, 2000.

[22] Robin Hanson, John Stutz, Peter Cheeseman
Bayesian Classification Theory
Technical Report FIA-90-12-7-01,NASA.

[23] R. Kohavi and F. Provost
Traffic incident detection: Sensors and algorithms
Machine Learning, vol. 39, no. 2-3, pp. 271-274, 1998.